

02-23-00

12

02/22/00

+

PTO/SB/05 (4/98)  
Approved for use through 09/30/2000. OMB 0651-0032  
Patent and Trademark Office: U.S. DEPARTMENT OF COMMERCE  
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

# UTILITY PATENT APPLICATION TRANSMITTAL

(Only for new nonprovisional applications under 37 C.F.R. § 1.53(b))

Attorney Docket No.	2459-1-003
First Inventor or Application Identifier	Ming-Ming Zhou
Title	METHODS OF IDENTIFYING ...
Express Mail Label No.	EL485954545US

## APPLICATION ELEMENTS

See MPEP chapter 600 concerning utility patent application contents.

ADDRESS TO: Assistant Commissioner for Patents  
Box Patent Application  
Washington, DC 20231

- ☐ \* Fee Transmittal Form (e.g., PTO/SB/17)  
(Submit an original and a duplicate for fee processing)
- ☒ Specification [Total Pages 67]  
(preferred arrangement set forth below)
  - Descriptive title of the Invention
  - Cross References to Related Applications
  - Statement Regarding Fed sponsored R & D
  - Reference to Microfiche Appendix
  - Background of the Invention
  - Brief Summary of the Invention
  - Brief Description of the Drawings (if filed)
  - Detailed Description
  - Claim(s)
  - Abstract of the Disclosure
- ☒ Drawing(s) (35 U.S.C. 113) [Total Sheets 9]
- Oath or Declaration unexecuted [Total Pages 3]
  - ☐ Newly executed (original or copy)
  - ☐ Copy from a prior application (37 C.F.R. § 1.63(d))  
(for continuation/divisional with Box 16 completed)
    - ☐ DELETION OF INVENTOR(S)  
Signed statement attached deleting inventor(s) named in the prior application, see 37 C.F.R. §§ 1.63(d)(2) and 1.33(b).

- ☐ Microfiche Computer Program (Appendix)
- Nucleotide and/or Amino Acid Sequence Submission (if applicable, all necessary)
  - ☒ Computer Readable Copy
  - ☒ Paper Copy (identical to computer copy)
  - ☒ Statement verifying identity of above copies

## ACCOMPANYING APPLICATION PARTS

- ☐ Assignment Papers (cover sheet & document(s))
- ☐ 37 C.F.R. § 3.73(b) Statement ☐ Power of Attorney  
(when there is an assignee)
- ☐ English Translation Document (if applicable)
- ☐ Information Disclosure Statement (IDS)/PTO-1449 ☐ Copies of IDS Citations
- ☐ Preliminary Amendment
- ☒ Return Receipt Postcard (MPEP 503)  
(Should be specifically itemized)
- ☐ \* Small Entity Statement filed in prior application, Status still proper and desired  
(PTO/SB/09-12)
- ☐ Certified Copy of Priority Document(s)  
(if foreign priority is claimed)
- ☒ Other: atomic coordinates  
in 6 Tables

\* NOTE FOR ITEMS 1 & 13: IN ORDER TO BE ENTITLED TO PAY SMALL ENTITY FEES, A SMALL ENTITY STATEMENT IS REQUIRED (37 C.F.R. § 1.27), EXCEPT IF ONE FILED IN A PRIOR APPLICATION IS RELIED UPON (37 C.F.R. § 1.28).

16. If a CONTINUING APPLICATION, check appropriate box, and supply the requisite information below and in a preliminary amendment.  
☐ Continuation ☐ Divisional ☐ Continuation-in-part (CIP) of prior application No: \_\_\_\_\_  
Prior application information: Examiner \_\_\_\_\_ Group / Art Unit: \_\_\_\_\_

For CONTINUATION or DIVISIONAL APPS only: The entire disclosure of the prior application, from which an oath or declaration is supplied under Box 4b, is considered a part of the disclosure of the accompanying continuation or divisional application and is hereby incorporated by reference. The incorporation can only be relied upon when a portion has been inadvertently omitted from the submitted application parts.

## 17. CORRESPONDENCE ADDRESS

☐ Customer Number or Bar Code Label (Insert Customer No. or Attach bar code label here) or ☐ Correspondence address below

Name	David A. Jackson Klauber & Jackson				
Address	Continental Plaza 411 Hackensack Avenue				
City	Hackensack	State	New Jersey	Zip Code	07601
Country	U.S.A.	Telephone	1(201)487-5800	Fax	1(201)343-1684

Name (Print/Type)	Michael D. Davis	Registration No. (Attorney/Agent)	39,161
Signature	Michael D. Davis	Date	2/22/2000

Burden Hour Statement: This form is estimated to take 0.2 hours to complete. Time will vary depending upon the needs of the individual case. Any comments on the amount of time you are required to complete this form should be sent to the Chief Information Officer, Patent and Trademark Office, Washington, DC 20231. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Assistant Commissioner for Patents, Box Patent Application, Washington, DC 20231.

+

**METHODS OF IDENTIFYING MODULATORS OF BROMODOMAINS**FIELD OF THE INVENTION

5

The present invention provides the three-dimensional structure of a histone acetyltransferase bromodomain. The three-dimensional structural information is included in the invention. The present invention also identifies for the first time, that bromodomains can bind to an acetylated binding partners. The interaction between bromodomains and their binding partners play a crucial role in various cellular functions, including in the regulation/modulation of DNA transcription. Therefore, the present invention provides procedures for identifying agents that can modulate the interaction of bromodomains and their binding partners by high throughput drug screening and/or through the use of rational drug design based on the three-dimensional data provided herein.

BACKGROUND OF THE INVENTION

In recent years great strides have been made in the elucidation of the steps involved in intercellular and intracellular signaling. Indeed, the individual steps of the cascade of events involved in a number of cellular signal transduction processes have been determined. For example, intercellular signal transduction generally begins with an intercellular ligand binding the extracellular portion of a receptor of the plasma membrane. The bound receptor then either directly or indirectly initiates the activation of one or more cellular factors. An activated cellular factor may act as transcription factor by entering the nucleus to interact with its corresponding genomic response element, or alternatively, it may interact with other cellular factors depending on the complexity of the process. In either case, one or more transcription factors ultimately bind to one or more specific genomic response elements. This binding plays a crucial role in the up and/or down regulation of the transcription of the specific genes that are under the control of these genomic response elements. However, the process of re-organizing the chromatin of eukaryotic cells, which is a prerequisite for the binding of the transcription factor to the genomic response elements, has remained a mystery.

- Chromatin contains several highly conserved histone proteins including: H3, H4, H2A, H2B, and H1. These histone proteins package eukaryotic DNA into repeating nucleosomal units that are folded into higher-order chromatin fibers [Luger and
- 5 Richmond, *Curr. Opin. Genet. Dev.* **8**:140-146 (1998)]. A portion of the histone that comprises roughly a quarter of the protein protrudes from the chromatin surface, and is thereby sensitive to proteolytic enzymes [van Holde, in *Chromatin* (Rich, A., ed., Springer, New York ) pages111-148 (1988); Hect *et al.*, *Cell* **80**:583-592 (1995)].
- This portion of the histone is known as the “histone tail”. Histone tails tend to be free
- 10 for protein-protein interaction, and are also the portion of the histone most prone to post-translational modification. Such post-translational modification includes acetylation, phosphorylation, methylation, ubiquitination, and ADP-ribosylation [van Holde, in *Chromatin* (Rich, A., ed., Springer, New York ) pages111-148 (1988)].
- 15 Of all classes of proteins, histones are amongst the most susceptible to post-translational modification. Perhaps the best studied post-translational modification of histones is the acetylation of specific lysine residues [Grunstin, M., *Nature*, **389**:349-352 (1997)]. Indeed, acetylation of histone lysine residues has been suggested to play a pivotal role in chromatin remodeling and gene activation. Consistently,
- 20 distinct classes of enzymes, namely histone acetyltransferases (HATs) and histone deacetylases (HDACs), acetylate or de-acetylate specific histone lysine residues [Struhl, *Genes Dev.* **12**:599-606 (1998)].
- Nearly all known nuclear HATs contain an approximately 110 amino acid sequence
- 25 known as the bromodomain [Jeanmougin *et al.*, *Trends in Biochemical Sciences*, **22**:151-153 (1997)], a protein motif that was initially discovered in *Drosophila* brahma protein. Bromodomains are found in a large number of chromatin-associated proteins and have now been identified in approximately 40 proteins, often adjacent to other protein motifs [Jeanmougin *et al.*, *Trends in Biochemical Sciences*, **22**:151-153
- 30 (1997); Tamkun *et al.*, *Cell*, **68**:561-572 (1992); Hanes *et al.*, *Nucleic Acids Research*, **20**:2603 (1992)]. Proteins that contain a bromodomain often contain a second bromodomain. However, despite the wide occurrence of bromodomains and their

likely role in chromatin regulation, their three-dimensional structure and binding partners heretofore have remained unknown.

Therefore, there is a need to identify a binding partner for a bromodomain. In addition, there is a need to identify agonists or antagonists to the bromodomain-binding partner complex. Since a preferred method of drug-screening relies on structure based drug design, there is also a need to determine the three-dimensional structure of a bromodomain. In this case, once the three dimensional structure of bromodomain is determined, potential agonists and/or potential antagonists can be designed with the aid of computer modeling [Bugg *et al.*, *Scientific American*, **Dec.**:92-98 (1993); West *et al.*, *TIPS*, **16**:67-74 (1995); Dunbrack *et al.*, *Folding & Design*, **2**:27-42 (1997)]. However, heretofore the three-dimensional structure of the bromodomain has remained unknown. Therefore, there is a need for obtaining a form of the bromodomain that is amenable for NMR analysis and/or X-ray crystallographic analysis. Furthermore, there is a need for the determination of the three-dimensional structure of the bromodomain. Finally, there is a need for procedures for related structural based drug design predicated on such structural data.

The citation of any reference herein should not be construed as an admission that such reference is available as "Prior Art" to the instant application.

### SUMMARY OF THE INVENTION

The present invention provides, for the first time, that bromodomains bind to acetyl-lysine residues of proteins. The present invention also provides the three-dimensional structure of a bromodomain as well as the three-dimensional structure of a bromodomain-acetyl-histamine complex. The structural information provided can be employed in methods of identifying drugs that can modulate the cellular processes that involve bromodomain-acetyl-lysine interactions. These interactions include chromatin remodeling, which is a required step in eukaryotic transcription. In a particular embodiment, the three-dimensional structural information is used in the design of an inhibitor of leukemia.



5

The present invention further provides a recombinant DNA molecule that comprises an isolated nucleic acid of the present invention, as described above, with or without a heterologous nucleotide sequence. Such a recombinant DNA molecule can be operatively linked to an expression control sequence and can be part of an expression vector. The present invention further provides a cell that comprises such an expression vector. The cell can be either a eukaryotic or a prokaryotic cell. The present invention further provides a method of expressing the peptides of the present invention or fragments thereof in this cell. One such method comprises culturing the cell in an appropriate cell culture medium under conditions that provide for expression of the peptide by the cell.

30 The present invention further provides a peptide consisting of about 21 to 40 amino acids that comprises a ZA loop of a bromodomain. In a preferred embodiment the

peptide comprises about 23 to 34 amino acids. The present invention also provides fusion proteins or peptides comprising these peptides.

In a preferred embodiment the peptide comprises the amino acid sequence of SEQ ID NO:3. In another embodiment the peptide comprises the amino acid sequence of SEQ ID NO:43. In particular embodiments the ZA loop is obtained from the bromodomain having the amino acid sequence of SEQ ID NO:7, or SEQ ID NO:8, or SEQ ID NO:9, or SEQ ID NO:10, or SEQ ID NO:11, or SEQ ID NO:12, or SEQ ID NO:13, or SEQ ID NO:14, or SEQ ID NO:15, or SEQ ID NO:16, or SEQ ID NO:17, or SEQ ID NO:18, or SEQ ID NO:19, or SEQ ID NO:20, or SEQ ID NO:21, or SEQ ID NO: 22, or SEQ ID NO:23, or SEQ ID NO:24, or SEQ ID NO:25, or SEQ ID NO:26, or SEQ ID NO:27, or SEQ ID NO:28, or SEQ ID NO:29, or SEQ ID NO:30, or SEQ ID NO: or SEQ ID NO:31, or SEQ ID NO:32, or SEQ ID NO: 33, or SEQ ID NO:34, or SEQ ID NO:35, or SEQ ID NO:36 , or SEQ ID NO:37, or SEQ ID NO:38, or SEQ ID NO: or SEQ ID NO:39, or SEQ ID NO:40, or SEQ ID NO:41, or SEQ ID NO:42.

The present invention also provides antibodies raised against the peptides/proteins of the present invention, or raised against an antigenic fragment of these proteins/fragments. In a particular embodiment an antibody is raised against a fragment of the ZA loop of a bromodomain. In another embodiment an antibody is raised against a fragment of a protein or peptide that comprises an acetyl-lysine, wherein the protein or peptide can bind to a bromodomain. Such fragments can be conjugated to a carrier protein or be part of a fusion protein. In one embodiment the antibody is a polyclonal antibody. In another embodiment, the antibody is a monoclonal antibody. A hybridoma that makes the monoclonal antibody is also part of the present invention. In a particular embodiment the antibody is a chimeric antibody. Antibodies that can specifically recognize acetyl-lysine residues involved bromodomain binding are also part of the present invention.

In another aspect of the present invention a method is provided for identifying a compound that modulates the affinity of a bromodomain for a ligand (and/or protein) that comprises an acetylated lysine. One such embodiment comprises contacting the

bromodomain and the ligand in the presence of a compound under conditions that ,  
the bromodomain and the ligand bind in the absence of the compound. The affinity of  
the bromodomain for the ligand is then determined (*e.g.*, measured). A compound is  
identified as a compound that modulates the affinity of the bromodomain for the  
5 ligand when there is a change in the affinity of the bromodomain for the ligand in the  
presence of the compound. When the affinity of the bromodomain for the ligand  
increases in the presence of the compound, the compound is identified as a promoting  
agent for the bromodomain-ligand complex. When the affinity of the bromodomain  
for the ligand decreases in the presence of the compound, the compound is identified  
10 as an inhibitor of the bromodomain-ligand complex. In a preferred embodiment, the  
compound to be tested is pre-selected by performing rational drug design with the set  
of atomic coordinates obtained from one or more of Tables 1-6. More preferably the  
selecting is performed in conjunction with computer modeling. In a particular  
embodiment, the compound is selected by performing rational drug design with the  
15 set of atomic coordinates obtained from a set of atomic coordinates defining the three-  
dimensional structure of a bromodomain consisting of the amino acid sequence of  
SEQ ID NO:7 alone or with acetyl-histamine.

The present invention also provides a method of identifying a compound that  
20 modulates the stability of a bromodomain-acetyl-lysine binding complex. One such  
embodiment comprises contacting the bromodomain-acetyl-lysine binding complex in  
the presence of the compound under conditions in which the bromodomain-acetyl-  
lysine binding complex forms in the absence of the compound. The stability of the  
bromodomain-acetyl-lysine binding complex is then determined (*e.g.*, measured). A  
25 compound is identified as a compound that modulates the stability of the  
bromodomain-acetyl-lysine binding complex, when there is a change in the stability  
of the bromodomain-acetyl-lysine binding complex in the presence of that compound.  
When the stability of the bromodomain-acetyl-lysine binding complex increases in the  
presence of the compound, the compound is identified as a stabilizing agent. When  
30 the stability of the bromodomain-acetyl-lysine binding complex decreases in the  
presence of the compound, the compound is identified as an inhibitor. In a preferred  
embodiment, the compound to be tested is pre-selected by performing rational drug

design with the set of atomic coordinates obtained from one or more of Tables 1-6. More preferably the selecting is performed in conjunction with computer modeling. In a particular embodiment, the compound is selected by performing rational drug design with the set of atomic coordinates obtained from a set of atomic coordinates  
 5 defining the three-dimensional structure of a bromodomain consisting of the amino acid sequence of SEQ ID NO:7 alone or with acetyl-histamine.

As anyone having skill in the art of drug development would readily understand, the potential drugs selected by the above methodologies can be refined by re-testing in  
 10 appropriate drug assays, including those disclosed herein. Chemical analogs of such potential drugs can be obtained (either through chemical synthesis or drug libraries) and be analogously tested. Therefore, methods comprising successive iterations of the steps of the individual drug assays, as exemplified herein, using either repetitive or different binding studies, or transcription activation studies or other such studies are  
 15 envisioned in the present invention. In addition, potential drugs may be identified first by rapid throughput drug screening, as described below, prior to performing computer modeling on a potential drug using the three-dimensional structure of the bromodomain.

20 The present invention further comprises all of the potential, selected, and putative compounds (drugs) identified by the methods of the present invention, as well as the final drugs themselves identified with the methods of the present invention.

The present invention further provides a method for identifying a potential binding  
 25 partner for a protein (*e.g.*, a histone) comprising an acetyl-lysine. One such embodiment comprises contacting the protein with a polypeptide comprising a bromodomain. In a preferred embodiment the bromodomain comprises the amino acid sequence of SEQ ID NO:3. In particular embodiments the bromodomain has the amino acid sequence of SEQ ID NO:7, or SEQ ID NO:8, or SEQ ID NO:9, or SEQ ID  
 30 NO:10, or SEQ ID NO:11, or SEQ ID NO:12, or SEQ ID NO:13, or SEQ ID NO:14, or SEQ ID NO:15, or SEQ ID NO:16, or SEQ ID NO:17, or SEQ ID NO:18, or SEQ ID NO:19, or SEQ ID NO:20, or SEQ ID NO:21, or SEQ ID NO: 22, or SEQ ID

NO:23, or SEQ ID NO:24, or SEQ ID NO:25, or SEQ ID NO:26, or SEQ ID NO:27,  
 or SEQ ID NO:28, or SEQ ID NO:29, or SEQ ID NO:30, or SEQ ID NO: or SEQ ID  
 NO:31, or SEQ ID NO:32, or SEQ ID NO: 33, or SEQ ID NO:34, or SEQ ID NO:35,  
 or SEQ ID NO:36 , or SEQ ID NO:37, or SEQ ID NO:38, or SEQ ID NO: or SEQ ID  
 5 NO:39, or SEQ ID NO:40, or SEQ ID NO:41, or SEQ ID NO:42.

The present invention further provides a method for identifying a protein having a  
 bromodomain. One such embodiment comprises contacting a cellular extract with a  
 peptide comprising an acetyl-lysine.

10

The present invention further provides agents that can inhibit the binding of a  
 bromodomain with a protein comprising an acetyl-lysine. In one embodiment the  
 agent is ISYGR-*AcK*-KRRQRR (SEQ ID NO:4). In another embodiment the agent is  
 ARKSTGG-*AcK*-APRKQL (SEQ ID NO:5). In still another embodiment the agent  
 15 is QSTSRHK-*AcK*-LMFKTE (SEQ ID NO:6). In yet another embodiment the agent  
 is an analog of acetyl-lysine such as acetyl-histamine. In still another embodiment the  
 agent is an antibody that recognizes an acetyl-lysine of a protein binding partner of a  
 bromodomain. In a preferred embodiment the agent is an antibody raised against a  
 ZA loop of a bromodomain. These agents can be used as pharmaceuticals in  
 20 compositions that contain a pharmaceutically acceptable carrier for example, or in the  
 various drug assays of the present invention, serving as controls to demonstrate  
 specificity.

Accordingly, it is a principal object of the present invention to provide the three-  
 25 dimensional coordinates of a bromodomain.

It is a further object of the present invention to provide the three-dimensional  
 coordinates of a bromodomain complexed with acetyl-histamine.

30 It is a further object of the present invention to provide an assay for identifying  
 proteins that contain bromodomains that bind proteins that comprise acetyl-lysine.

00510314.022200

It is a further object of the present invention to provide methods of identifying drugs that can inhibit the binding of a bromodomain to a protein containing acetyl-lysine.

10 It is a further object of the present invention to provide a method of identifying drugs  
that can treat leukemia.

15 These and other aspects of the present invention will be better appreciated by  
reference to the following drawings and Detailed Description.

Figure 1. Structure-based sequence alignment of a selected number of bromodomains. The sequences were aligned based on the NMR-derived structure of the P/CAF bromodomain, and the predicated four  $\alpha$ -helices are shown in green boxes. Bromodomains are grouped on the basis of the sequence and/or functional similarities as described by Jeanmougin *et al.*, [Trends in Biochemical Sciences, **22**:151-153 (1997)]. Residue numbers of the P/CAF bromodomain are indicated above its sequence. Three absolutely conserved residues, corresponding to Pro751, Pro767, and Asn803 in the P/CAF bromodomain, are shown in red. Highly conserved residues are colored in blue. The residues of the P/CAF bromodomain that interact with acetyl-histamine, as determined by intermolecular NOEs, are indicated by asterisks. The ZA loop, which is critical for acetyl-lysine binding, for each of the indicated bromodomains is also identified. The underlined residues were changed individually

by site-directed mutagenesis to Ala. Genbank accession numbers for the proteins are as indicated in Table 8, in the Example below, along with the SEQ ID NOs. for the bromodomain sequences.

- 5 Figures 2A-2H depict the structure of the P/CAF bromodomain. Figures 2A-2B shows the stereoview of the  $C_\alpha$  trace of 30 superimposed NMR-derived structures of the bromodomain (residues 722-830). The N-terminal four residues (SKEP) which are structurally disordered are omitted for clarity. For the final 30 structures, the root-mean-square deviations (RMSDs) of the backbone and all heavy atoms are  $0.63 \pm 0.11 \text{ \AA}$  and  $1.15 \pm 0.12 \text{ \AA}$  for residues 723-830, respectively. The RMSDs of the backbone and all heavy atoms for the four  $\alpha$ -helices (residues 727-743, 770-776, 785-802, and 807-827), are  $0.34 \pm 0.04 \text{ \AA}$  and  $0.87 \pm 0.06 \text{ \AA}$ , respectively. Figures 2C-2D show the stereoview of the bromodomain structures from the bottom of the protein, which is rotated approximately  $90^\circ$  from the orientation in Figures 2A-2B.
- 10  $\pm 0.11 \text{ \AA}$  and  $1.15 \pm 0.12 \text{ \AA}$  for residues 723-830, respectively. The RMSDs of the backbone and all heavy atoms for the four  $\alpha$ -helices (residues 727-743, 770-776, 785-802, and 807-827), are  $0.34 \pm 0.04 \text{ \AA}$  and  $0.87 \pm 0.06 \text{ \AA}$ , respectively. Figures 2C-2D show the stereoview of the bromodomain structures from the bottom of the protein, which is rotated approximately  $90^\circ$  from the orientation in Figures 2A-2B.
- 15 Figure 2E shows the Ribbons [Carson, M., *J. Appl. Crystallogr.* **24**:958-961 (1991)] depiction of the averaged minimized NMR structure of the P/CAF bromodomain. The orientation of Figure 2E is as shown in Figures 2A-2B. Figures 2F-2G are schematic representations of the overall topology of the up-and-down four-helix bundle folds with the opposite handedness. The left-handed fold is seen in bromodomain, cytochrome  $b_5$ , and T4 lysozyme (left, Figure 2F), whereas the right-handed four-helix bundles are observed in proteins such as hemerythrin and cytochrome  $b_{562}$  (right, Figure 2G) [Richardson, J., *Adv. Protein Chem.*, **34**:167-339 (1989); Presnell and Cohen, *Proc. Natl. Acad. Sci. USA* **86**:6592-6596 (1989)].
- 20 bromodomain, cytochrome  $b_5$ , and T4 lysozyme (left, Figure 2F), whereas the right-handed four-helix bundles are observed in proteins such as hemerythrin and cytochrome  $b_{562}$  (right, Figure 2G) [Richardson, J., *Adv. Protein Chem.*, **34**:167-339 (1989); Presnell and Cohen, *Proc. Natl. Acad. Sci. USA* **86**:6592-6596 (1989)].
- Figure 2H is a molecular surface representation of the electrostatic potential (blue = positive; red = negative) of the bromodomain calculated in GRASP [Nicholls *et al.*, *Biophys. J.* **64**:166-170 (1993)]. The hydrophobic and aromatic residues (Tyr809, Tyr802, Tyr760, Ala757, and Val752) located between the ZA and BC loops are indicated.
- 25 positive; red = negative) of the bromodomain calculated in GRASP [Nicholls *et al.*, *Biophys. J.* **64**:166-170 (1993)]. The hydrophobic and aromatic residues (Tyr809, Tyr802, Tyr760, Ala757, and Val752) located between the ZA and BC loops are indicated.
- 30 Figures 3A-3C show the binding of the P/CAF bromodomain to AcK. Figure 3A shows the superimposed region of the 2D  $^{15}\text{N}$ -HSQC spectra of the bromodomain (approximately 0.5 mM) in its free form (red) and complexed to the AcK-containing

H4 peptide (molar ratio 1:6) (black). Figure 3B is the Ribbon and dotted-surface diagram of the bromodomain depicting the location of the lysine-acetylated H4 peptide binding site. The color coding reflects the chemical shift changes ( $\Delta\delta$ ) of the backbone amide  $^1\text{H}$  and  $^{15}\text{N}$  resonances upon binding to the AcK peptide as observed in the  $^{15}\text{N}$ -HSQC spectra. The normalized weighted average of the chemical shift changes was calculated by  $\Delta_{av}/\Delta_{max} = [\Delta\delta_{\text{NH}}^2 + \Delta\delta_{\text{N}}^2/25]/2]^{1/2}/\Delta_{max}$ , where  $\Delta_{max}$  is the maximum weighted chemical shift difference observed for Tyr809 (0.16ppm). The backbone atoms are color-coded in red, yellow, or green for residues that have  $\Delta_{av}/\Delta_{max}$  of  $>0.6$  (Tyr809, Glu808, Asn803, and Ala757),  $0.2-0.6$  (Ala813, Tyr802, Tyr760, and Val752), or  $<0.2$  (Cys812, Ser807, Cys799, Phe796, and Phe748), respectively. The non-perturbed residues are shown in blue. Figure 3C shows the chemical structures of acetyl-lysine, acetyl-histamine, and acetyl-histidine.

Figure 4 depicts the acetyl-lysine binding pocket. This is the Ribbons [Carson, M., *J. Appl. Crystallogr.* **24**:958-961 (1991)] depiction of a portion of the P/CAF bromodomain complexed with the acetyl-histamine. The ligand is color-coded by atom type.

## DETAILED DESCRIPTION OF THE INVENTION

The present invention identifies a general binding partner (ligand) for the protein motif known as the bromodomain. Indeed, by combining structural and site-directed mutagenesis studies the present invention demonstrates that bromodomains can interact specifically with acetyl-lysine (AcK), making them the first protein modules known to exhibit such interactions. Like other modular domains, such as Src homology-2 (SH2) and phosphotyrosine binding (PTB) domains, which specifically interact with phosphotyrosine-containing proteins, the bromodomain/acetyl-lysine recognition provides a means to regulate protein-protein interactions via protein lysine acetylation. The nature of the acetyl-lysine recognition by the bromodomain is similar to that of histone acetyltransferase interaction with acetyl-CoA. The present invention therefore couples for the first time, the functionality of the bromodomain with the HAT activity of coactivators in the regulation of gene transcription.



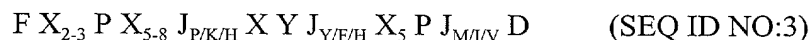
The present invention further provides both a nuclear magnetic resonance (NMR) structure of the bromodomain from the HAT coactivator P/CAF (p300/CBP-associated factor) as well as the structure for the P/CAF bromodomain in complex with acetyl-histamine. The structure reveals an unusual left-handed up-and-down four-helix bundle.

The results disclosed herein explain prior deletion experiments which showed that the bromodomain is indispensable for the function of GCN5 in yeast.

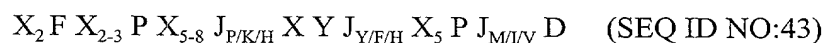
Bromodomain-AcK binding also appears to be important for the assembly and activity of multiprotein complexes in transcriptional activation. The results reported herein therefore, form the foundation for identifying specific biological ligands and for defining the molecular mechanisms by which the extensive family of bromodomains participate in chromatin remodeling and transcriptional activation

As disclosed herein, the binding partner for the bromodomain is a peptide or protein comprising an acetyl-lysine (AcK). Interestingly, whereas a free acetyl-lysine does not appear to bind the bromodomain, an analog of the acetyl-lysine, acetyl-histamine, does. This is most likely due to the additional charge present in the free amino acid. Consistently, free acetyl-histidine also does not to bind the bromodomain.

The present invention further provides a key region of the bromodomain for the interaction with its acetyl-lysine binding partner, the ZA loop. The amino acid sequence of the ZA loop is defined in Figure 1 for a number of bromodomains and is depicted in Figure 2A for P/CAF. In a particular embodiment, the ZA loop has between about 21 and 40 amino acid residues comprising the amino acid sequence of :



more preferably the ZA loop has about 23 to 34 amino acid residues and comprises the amino acid sequence:



(1) The single letter amino acid code is used in this description, *i.e.*, “F” for phenylalanine; “P” for proline; “Y” for tyrosine; and “D” for aspartic acid.

(2) “X” indicates any amino acid (an undesignated amino acid); and X, X<sub>2</sub>, X<sub>2-3</sub>, X<sub>5</sub>, and X<sub>5-8</sub> indicates one undesignated amino acid, two consecutive undesignated amino acids, two or three consecutive undesignated amino acids, five consecutive undesignated amino acids, and five to eight consecutive undesignated amino acids respectively.

(3) “J” indicates that identity of the amino acid is restricted to a particular group, again the one letter code is used

- 10 : (i) J<sub>P/K/H</sub> is either proline, lysine or histidine.  
 (ii) J<sub>Y/F/H</sub> is either tyrosine, phenylalanine or histidine.  
 (iii) J<sub>M/I/V</sub> is either methionine, isoleucine, or valine.

Since this region of the bromodomain is important in binding its acetyl-lysine binding partner, antibodies specifically raised against this region are also included in the present invention. In a particular embodiment, the antibody is a humanized chimeric antibody that can be used in therapeutic treatment. Thus monoclonal, chimeric, and polyclonal antibodies raised against bromodomains, preferably against amino acid residues in the ZA loop region are part of the present invention. In a specific embodiment the antibody is raised against a peptide, fusion peptide or conjugated peptide consisting of amino acid residues 746 to 765 of SEQ ID NO:2, *i.e.*, WPFMEPVKRTEAPGYEYEVIR (SEQ ID NO:44). Such antibodies can be used in the treatment of leukemia for example. Alternatively, these antibodies can be used in drug discovery assays.

25 Thus the present invention provides the first detailed structural information regarding a bromodomain and a bromodomain complexed with its acetylated binding partner. The present invention therefore provides the three-dimensional structure of the bromodomain and a bromodomain acetylated binding partner complex. Since the interaction of the bromodomain with a histone for example, can play a significant role in chromatin remodeling/regulation, the structural information provided herein can be employed in methods of identifying drugs that can modulate basic cell processes by modulating the transcription. In a particular embodiment, the three-dimensional

structural information is used in the design of a small organic molecule for the treatment of cancer.

Indeed, the bromodomain and lysine-acetylated protein interaction can now be  
 5 implicated to play a causal role in the development of a number of diseases including  
 cancers such as leukemia. For example, chromatin remodeling plays a central role in  
 the etiology of viral infection and cancer [Archer and Hodin, *Curr. Opin. Genet. Biol.*  
*9*:171-174 (1999); Jacobson and Pillus, *Curr. Opin. Genet. Biol.* *9*:175-184 (1999)].  
 Both altered histone acetylation/deacetylation and aberrant forms of chromatin-  
 10 remodeling complexes are associated with human diseases. Furthermore,  
 chromosomal translocation of various cellular genes with those encoding HATs and  
 subunits of chromatin remodeling complexes have been implicated in leukomogenesis.  
 The *MOZ* (monocytic leukemia zinc finger) and *MLL/ALL-1* genes are frequently fused  
 to the gene encoding the co-activator HAT CBP [Sobulo *et al.*, *Proc. Natl. Acad. Sci.*  
 15 *USA* *94*:8732-8737(1997)]. The resulting fusion protein MLL-CBP contains the  
 tandem bromodomain-PHD finger-HAT domain of CBP. It also has been shown that  
 both the bromodomain and HAT domain of CBP are required for leukomogenesis,  
 because deletion of either the bromodomain or the HAT domain results in loss of the  
 MLL-CBP fusion protein's ability for cell transform. These results indicate that the  
 20 CBP bromodomain, and more particularly, the ZA loop of the CBP bromodomain, is  
 an excellent target for developing drugs that interfere with the bromodomain acetyl-  
 lysine interaction that can be used in the treatment of human acute leukemia. In  
 addition, an antibody (*e.g.*, a humanized antibody) raised specifically against a peptide  
 from the ZA loop of the CBP bromodomain could also be effective for treating these  
 25 conditions.

Furthermore, the human immunodeficiency virus type 1 (HIV-1) *trans*-activator  
 protein, Tat, is absolutely required for productive HIV viral replication [Jeang and  
 Gatignol, *Curr. Top. Microbiol. Immunol.*, *188*:123-144(1994)]. Recently, it has been  
 30 shown that HIV-1 Tat transcriptional activity is tightly regulated by lysine acetylation  
 [Kiernan *et al.*, *EMBO Journal* *18*:6106-6118 (1999)]. Therefore, the interaction of  
 the acetyl-lysine of Tat with one or more bromodomain-containing proteins associated

with chromatin remodeling could mediate gene transcription. Thus, the bromodomain/lysine-acetylated Tat interaction could also serve as a drug target for blocking HIV replication in cells. Similarly, an antibody raised specifically against a peptide from the ZA loop of the bromodomain could also be effective for treating these conditions.

In addition, based on the new structural information disclosed herein, the key amino acid residues for the binding of a given bromodomain and its binding partner can be identified and further elucidated using basic mutagenesis and standard isothermal titration calorimetry, for example. In this case, both the crucial amino acids for the bromodomain and the binding partner (i.e., apart from the acetyl-lysine) can be readily determined and are also part of the present invention.

The results obtained from the structural and functional studies disclosed herein provide the foundation for both high throughput drug screening and structure-based rational drug design. The agents identified by this procedure will be useful for ameliorating conditions involving chromatin remodeling/regulation as indicated above.

Structure based rational drug design is the most efficient method of drug development. However, heretofore, no information has been disclosed regarding the structure of the bromodomain or more importantly, its interaction with the acetyl-lysine of its binding partner. Obtaining detailed structural information requires an extensive NMR or X-ray crystallographic analysis. By determining and then exploiting the detailed structural information of the bromodomain and of the bromodomain/acetyl-histamine (exemplified by NMR analysis below) the present invention provides novel methods for developing new drugs through structure based rational drug design.

Thus the present invention provides representative sets of the atomic structure coordinates of the free form of the P/CAF bromodomain (Table 5) and of the P/CAF bromodomain-acetyl-histamine complex (Table 6) which were both obtained by NMR analysis. A Ribbon diagram of the three-dimensional structure of the P/CAF bromodomain is depicted in Figure 2E, whereas the P/CAF bromodomain acetyl-lysine

binding pocket is depicted in Figure 4. The present invention also provides the NOE-derived distance restraints, and NMR chemical shift assignments of the P/CAF bromodomain. The NMR chemical shift assignments of the P/CAF bromodomain are included in the chemical shift table (Table 1) for the  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum of P/CAF bromodomain. The unambiguous NOE-derived Inter-proton Distance Restraints (Table 2), the ambiguous NOE-derived Inter-proton Distance Restraints (Table 3) and the  $^1\text{H}$  bonding restraints (Table 4) are also disclosed herein. The sample atomic coordinate data provided enable the skilled artisan to practice the invention. In addition, Tables 1-6 are also capable of being placed into a computer readable form which is also part of the present invention. Furthermore, methods of using these coordinates and chemical shifts and related information (including in computer readable forms) either individually or together in drug assays are also provided. More particularly, such atomic coordinates can be used to identify potential ligands or drugs which will modulate the binding of a bromodomain with its binding partner.

Therefore, if appearing herein, the following terms shall have the definitions set out below.

As used herein a “bromodomain-acetyl-lysine binding complex” is a binding complex between a bromodomain or fragment thereof and either a peptide/polypeptide comprising an acetyl-lysine (or an analog of acetyl-lysine), or a free analog of acetyl-lysine, such as acetyl-histamine disclosed in the Example below. Preferably, the peptide comprises at least six amino acids in addition to the acetyl-lysine. The dissociation constant of a bromodomain-acetyl-lysine binding complex is dependent on whether the lysine residue or analog thereof is acetylated or not, such that the affinity for the bromodomain and the peptide comprising the lysine residue (for example) significantly decreases when that lysine residue is not acetylated.

As used herein a “ZA loop” of a bromodomain is one portion of a bromodomain that is involved in the binding of the bromodomain to the acetyl-lysine. The structure of the ZA loop of the bromodomain of for P/CAF is depicted in Figure 2A. The ZA loop has between about 20 and 40 amino acids and comprises the amino acid sequence of SEQ ID NO:3. More preferably the ZA loop comprises between about 23 to 34 amino acids

and has the amino acid sequence SEQ ID NO:43. The amino acid sequence of the ZA loop for a representative number of individual bromodomains is shown in Figure 1.

A "polypeptide" or "peptide" comprising a fragment of a bromodomain, such as the  
5 ZA loop, or a peptide or polypeptide comprising an acetyl-lysine, as used herein can be the "fragment" alone, or a larger chimeric or fusion peptide/protein which contains the "fragment".

As used herein the terms "fusion protein" and "fusion peptide" are used  
10 interchangeably and encompass "chimeric proteins and/or chimeric peptides" and fusion "intein proteins/peptides". A fusion protein comprises at least a portion of a protein or peptide of the present invention, *e.g.*, a bromodomain, joined *via* a peptide bond to at least a portion of another protein or peptide including *e.g.*, a second bromodomain in a chimeric fusion protein. In a particular embodiment the portion of  
15 the bromodomain is antigenic. Fusion proteins can comprise a marker protein or peptide, or a protein or peptide that aids in the isolation and/or purification of the protein, for example.

As used herein, and unless otherwise specified, the terms "agent", "potential drug",  
20 "compound", "test compound" or "potential compound" are used interchangeably, and refer to chemicals which potentially have a use as an inhibitor or activator/stabilizer of bromodomain-acetyl-lysine binding. Therefore, such "agents", "potential drugs", "compounds" and "potential compounds" may be used, as described herein, in drug assays and drug screens and the like.

25 As used herein a "small organic molecule" is an organic compound, including a peptide [or organic compound complexed with an inorganic compound (*e.g.*, metal)] that has a molecular weight of less than 3 Kilodaltons. Such small organic molecules can be included as agents, etc. as defined above.

30 As used herein the term "binds to" is meant to include all such specific interactions that result in two or more molecules showing a preference for one another relative to some third molecule. This includes processes such as covalent, ionic, hydrophobic and

hydrogen bonding but does not include non-specific associations such as solvent preferences.

As used herein the term "about" signifies that a value is within twenty percent of the indicated value *i.e.*, a peptide containing "about" 20 amino acid residues can contain between 16 and 24 amino acid residues.

General Techniques for Constructing Nucleic Acids That Encode the Bromodomains and Fragments Thereof (Including, ZA Loops); and the Bromodomain Binding

10 Partners of the Present Invention.

In accordance with the present invention there may be employed conventional molecular biology, microbiology, and recombinant DNA techniques within the skill of the art. Such techniques are explained fully in the literature. See, *e.g.*, Sambrook, 15 Fritsch & Maniatis, *Molecular Cloning: A Laboratory Manual*, Second Edition (1989) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (herein "Sambrook *et al.*, 1989"); *DNA Cloning: A Practical Approach*, Volumes I and II (D.N. Glover ed. 1985); *Oligonucleotide Synthesis* (M.J. Gait ed. 1984); *Nucleic Acid Hybridization* [B.D. Hames & S.J. Higgins eds. (1985)]; *Transcription And* 20 *Translation* [B.D. Hames & S.J. Higgins, eds. (1984)]; *Animal Cell Culture* [R.I. Freshney, ed. (1986)]; *Immobilized Cells And Enzymes* [IRL Press, (1986)]; B. Perbal, *A Practical Guide To Molecular Cloning* (1984); F.M. Ausubel *et al.* (eds.), *Current Protocols in Molecular Biology*, John Wiley & Sons, Inc. (1994).

25 Therefore, if appearing herein, the following terms shall have the definitions set out below.

As used herein, the term "gene" refers to an assembly of nucleotides that encode a polypeptide, and includes cDNA and genomic DNA nucleic acids.

30

A "vector" is a replicon, such as plasmid, phage or cosmid, to which another DNA segment may be attached so as to bring about the replication of the attached segment. A "replicon" is any genetic element (*e.g.*, plasmid, chromosome, virus) that functions

as an autonomous unit of DNA replication *in vivo*, *i.e.*, capable of replication under its own control.

A "cassette" refers to a segment of DNA that can be inserted into a vector at specific  
 5 restriction sites. The segment of DNA encodes a polypeptide of interest, and the  
 cassette and restriction sites are designed to ensure insertion of the cassette in the  
 proper reading frame for transcription and translation.

A cell has been "transfected" by exogenous or heterologous DNA when such DNA has  
 10 been introduced inside the cell.

A "nucleic acid molecule" refers to the phosphate ester polymeric form of  
 ribonucleosides (adenosine, guanosine, uridine or cytidine; "RNA molecules") or  
 deoxyribonucleosides (deoxyadenosine, deoxyguanosine, deoxythymidine, or  
 15 deoxycytidine; "DNA molecules"), or any phosphoester analogues thereof, such as  
 phosphorothioates and thioesters, in either single stranded form, or a double-stranded  
 helix. Double stranded DNA-DNA, DNA-RNA and RNA-RNA helices are possible.  
 The term nucleic acid molecule, and in particular DNA or RNA molecule, refers only  
 to the primary and secondary structure of the molecule, and does not limit it to any  
 20 particular tertiary forms. Thus, this term includes double-stranded DNA found, *inter*  
*alia*, in linear or circular DNA molecules (*e.g.*, restriction fragments), plasmids, and  
 chromosomes. In discussing the structure of particular double-stranded DNA  
 molecules, sequences may be described herein according to the normal convention of  
 giving only the sequence in the 5' to 3' direction along the nontranscribed strand of  
 25 DNA (*i.e.*, the strand having a sequence homologous to the mRNA). A "recombinant  
 DNA molecule" is a DNA molecule that has undergone a molecular biological  
 manipulation.

A nucleic acid molecule is "hybridizable" to another nucleic acid molecule, such as a  
 30 cDNA, genomic DNA, or RNA, when a single stranded form of the nucleic acid  
 molecule can anneal to the other nucleic acid molecule under the appropriate  
 conditions of temperature and solution ionic strength (*see* Sambrook *et al.*, *supra*).  
 The conditions of temperature and ionic strength determine the "stringency" of the



hybridization. For preliminary screening for homologous nucleic acids, low stringency hybridization conditions, corresponding to a  $T_m$  of  $55^\circ$ , can be used, *e.g.*, 5x SSC, 0.1% SDS, 0.25% milk, and no formamide; or 30% formamide, 5x SSC, 0.5% SDS). Moderate stringency hybridization conditions correspond to a higher  $T_m$ , *e.g.*, 40% formamide, with 5x or 6x SCC. High stringency hybridization conditions correspond to the highest  $T_m$ , *e.g.*, 50% formamide, 5x or 6x SCC. Hybridization requires that the two nucleic acids contain complementary sequences, although depending on the stringency of the hybridization, mismatches between bases are possible. The appropriate stringency for hybridizing nucleic acids depends on the length of the nucleic acids and the degree of complementation, variables well known in the art. The greater the degree of similarity or homology between two nucleotide sequences, the greater the value of  $T_m$  for hybrids of nucleic acids having those sequences. The relative stability (corresponding to higher  $T_m$ ) of nucleic acid hybridizations decreases in the following order: RNA:RNA, DNA:RNA, DNA:DNA. For hybrids of greater than 100 nucleotides in length, equations for calculating  $T_m$  have been derived (*see* Sambrook *et al.*, *supra*, 9.50-10.51). For hybridization with shorter nucleic acids, *i.e.*, oligonucleotides, the position of mismatches becomes more important, and the length of the oligonucleotide determines its specificity (*see* Sambrook *et al.*, *supra*, 11.7-11.8). Preferably a minimum length for a hybridizable nucleic acid is at least about 12 nucleotides; preferably at least about 18 nucleotides; and more preferably the length is at least about 27 nucleotides; and most preferably 36 nucleotides.

In a specific embodiment, the term "standard hybridization conditions" refers to a  $T_m$  of  $55^\circ\text{C}$ , and utilizes conditions as set forth above. In a preferred embodiment, the  $T_m$  is  $60^\circ\text{C}$ ; in a more preferred embodiment, the  $T_m$  is  $65^\circ\text{C}$ .

A DNA "coding sequence" is a double-stranded DNA sequence which is transcribed and translated into a polypeptide in a cell *in vitro* or *in vivo* when placed under the control of appropriate regulatory sequences. The boundaries of the coding sequence are determined by a start codon at the 5' (amino) terminus and a translation stop codon at the 3' (carboxyl) terminus. A coding sequence can include, but is not limited to, prokaryotic sequences and synthetic DNA sequences. If the coding sequence is

intended for expression in a eukaryotic cell, a polyadenylation signal and transcription termination sequence will usually be located 3' to the coding sequence.

Transcriptional and translational control sequences are DNA regulatory sequences, such as promoters, enhancers, terminators, and the like, that provide for the expression of a coding sequence in a host cell. In eukaryotic cells, polyadenylation signals are control sequences.

A "promoter sequence" is a DNA regulatory region capable of binding RNA polymerase in a cell and initiating transcription of a downstream (3' direction) coding sequence. For purposes of defining the present invention, the promoter sequence is bounded at its 3' terminus by the transcription initiation site and extends upstream (5' direction) to include the minimum number of bases or elements necessary to initiate transcription at levels detectable above background. Within the promoter sequence will be found a transcription initiation site (conveniently defined for example, by mapping with nuclease S1), as well as protein binding domains (consensus sequences) responsible for the binding of RNA polymerase.

A coding sequence is "under the control" of transcriptional and translational control sequences in a cell when RNA polymerase transcribes the coding sequence into mRNA, which is then trans-RNA spliced and translated into the protein encoded by the coding sequence.

A DNA sequence is "operatively linked" to an expression control sequence when the expression control sequence controls and regulates the transcription and translation of that DNA sequence. The term "operatively linked" includes having an appropriate start signal (e.g., ATG) in front of the DNA sequence to be expressed and maintaining the correct reading frame to permit expression of the DNA sequence under the control of the expression control sequence and production of the desired product encoded by the DNA sequence. If a gene that one desires to insert into a recombinant DNA molecule does not contain an appropriate start signal, such a start signal can be inserted in front of the gene.

5

10

15

25

30

commercially available programs can also be used to determine sequence similarity using the same or analogous default parameters.

- The term "corresponding to" is used herein to refer similar or homologous sequences, whether the exact position is identical or different from the molecule to which the similarity or homology is measured. Thus, the term "corresponding to" refers to the sequence similarity, and not the numbering of the amino acid residues or nucleotide bases.
- As used herein a "heterologous nucleotide sequence" is a nucleotide sequence that is added to a nucleotide sequence of the present invention by recombinant methods to form a nucleic acid which is not naturally formed in nature. Such nucleic acids can encode fusion proteins or peptides, including chimeric proteins and peptides. Thus the heterologous nucleotide sequence can encode peptides and/or proteins which contain regulatory and/or structural properties. In another such embodiment the heterologous nucleotide can encode a protein or peptide that functions as a means of detecting the protein or peptide encoded by the nucleotide sequence of the present invention after the recombinant nucleic acid is expressed. In still another such embodiment the heterologous nucleotide can function as a means of detecting a nucleotide sequence of the present invention. A heterologous nucleotide sequence can comprise non-coding sequences including restriction sites, regulatory sites, promoters and the like.

- The present invention also relates to cloning vectors containing nucleic acids encoding analogs and derivatives of the bromodomains of the present invention and polypeptides/peptides that can bind a bromodomain when a lysine of the polypeptide/peptide is acetylated, including modified fragments, that have the same or homologous functional activity as the individual fragments, and homologs thereof. The production and use of derivatives and analogs related to the fragments are within the scope of the present invention.

- Due to the degeneracy of nucleotide coding sequences, other DNA sequences which encode substantially the same amino acid sequence as a nucleic acid encoding a protein

5

25

- 30

A conservative change generally leads to less change in the structure and function of the resulting protein. A non-conservative change is more likely to alter the structure,

activity or function of the resulting protein. The present invention should be considered to include sequences containing conservative changes which do not significantly alter the activity or binding characteristics of the resulting protein.

Specific amino acid residues for the P/CAF bromodomain have been identified that are  
 5 important for binding, indicating a potential lower stringency for the substitution of the remaining amino acids residues.

All of the peptides/fragments of the present invention can be modified by being placed in a fusion or chimeric peptide or protein, or labeled *e.g.*, to have an N-terminal FLAG-  
 10 tag, or H6 tag. In a particular embodiment the P/CAF bromodomain fragment can be modified to contain a marker protein such as green fluorescent protein as described in U.S. Patent No. 5,625,048 filed April 29, 1997 and WO 97/26333, published July 24, 1997 each of which are hereby incorporated by reference herein in their entireties.

15 The nucleic acids encoding peptides and protein fragments of the present invention and analogs thereof can be produced by various methods known in the art. The manipulations which result in their production can occur at the gene or protein level [Sambrook *et al.*, 1989, *supra*]. The nucleotide sequence can be cleaved at appropriate sites with restriction endonuclease(s), followed by further enzymatic modification if  
 20 desired, isolated, and ligated *in vitro*. In addition a nucleic acid sequence can be mutated *in vitro* or *in vivo*, to create and/or destroy translation, initiation, and/or termination sequences, or to create variations in coding regions and/or form new restriction endonuclease sites or destroy preexisting ones, to facilitate further *in vitro* modification. Any technique for mutagenesis known in the art can be used, including  
 25 but not limited to, *in vitro* site-directed mutagenesis [Hutchinson *et al.*, *J. Biol. Chem.*, **253**:6551 (1978); Zoller and Smith, *DNA*, **3**:479-488 (1984); Oliphant *et al.*, *Gene*, **44**:177 (1986); Hutchinson *et al.*, *Proc. Natl. Acad. Sci. U.S.A.*, **83**:710 (1986)], use of TAB® linkers (Pharmacia), etc. PCR techniques are preferred for site directed mutagenesis [see Higuchi, 1989, "Using PCR to Engineer DNA", in *PCR Technology: Principles and Applications for DNA Amplification*, H. Erlich, ed., Stockton Press,  
 30 Chapter 6, pp. 61-70].

5

A bacterial protein expression system can be used to make various stable isotopically labeled ( $^{13}\text{C}$ ,  $^{15}\text{N}$ , and  $^2\text{H}$ ) protein samples that are useful for a three-dimensional NMR structural determination of a protein complex. For example a pET14b (Novagen) bacterial expression vector can be constructed which expresses the recombinant P/CAF  
10 bromodomain as an amino-terminal His-tagged fusion protein.

Protein expression and purification can be conducted using standard procedures for His-tagged proteins [Zhou *et al.*, *J. Biol. Chem.* **270**:31119-31123 (1995)]. To optimize the level of protein expression, various bacterial growth and expression conditions can be screened, which include different *E. Coli* cell lines, and growth and protein induction temperatures. Generally, it is preferred to obtain the maximum amount of soluble protein while still inducing protein expression with a relatively low IPTG concentration *e.g.*, ~0.2mM (final concentration) at 16°C. As exemplified below, the bromodomain of P/CAF (residues 719-832 of SEQ ID NO:2 which is SEQ ID NO:7) was subcloned into the pET14b expression vector (Novagen) and expressed in *Escherichia coli* BL21(DE3) cells. Uniformly <sup>15</sup>N- and <sup>15</sup>N/<sup>13</sup>C-labeled proteins were prepared by growing bacteria in a minimal medium containing <sup>15</sup>NH<sub>4</sub>Cl with or without <sup>13</sup>C<sub>6</sub>-glucose. A uniformly <sup>15</sup>N/<sup>13</sup>C-labeled and fractionally deuterated protein sample was prepared by growing the cells in 75% <sup>2</sup>H<sub>2</sub>O. The bromodomain was purified by affinity chromatography on a nickel-IDA column (Invitrogen) followed by the removal of poly-His tag by thrombin cleavage. The final purification of the protein was achieved by size-exclusion chromatography. The acetyl-lysine-containing peptides were prepared on a MilliGen 9050 peptide synthesizer (Perkin Elmer) using Fmoc/HBTU chemistry. Acetyl-lysine was incorporated using the reagent Fmoc-Ac-Lys with HBTU/DIPEA activation. NMR samples contained approximately 1 mM protein in 100mM phosphate buffer of pH 6.5 and 5mM perdeuterated DTT and 0.5mM EDTA in H<sub>2</sub>O/<sup>2</sup>H<sub>2</sub>O (9/1) or <sup>2</sup>H<sub>2</sub>O.

One major advantage of using the heteronuclear multidimensional approach, as exemplified herein, is that the NMR resonance assignments of a protein are obtained in a sequence-specific manner which assures accuracy and greatly facilitates data analysis and structure determination [Clare, G. M. & Gronenborn, A. M. *Meth. Enzymol.*

- 5 **239:249-363 (1994)]**. In addition, the signal overlapping problems in the protein spectra are minimized by the use of multidimensional NMR spectra, which separates the proton signals according to the chemical shifts of their attached hetero-nuclei (such as  $^{15}\text{N}$  and  $^{13}\text{C}$ ). This NMR approach has been proven very powerful for structural analysis of large proteins [Clare, G. M. & Gronenborn, A. M. *Meth. Enzymol.*
- 10 **239:249-363 (1994)]**. To facilitate sequence-specific resonance assignments for the structural study, a uniformly  $^{13}\text{C}$ ,  $^{15}\text{N}$ -labeled and fractionally (75%) deuterated protein sample of the bromodomain can be prepared by growing bacterial cells in 75%  $^2\text{H}_2\text{O}$  as exemplified below. Such protein samples can be used for triple-resonance NMR experiments. A triple-labeled protein sample is useful for high-resolution NMR
- 15 structural studies. Because of the favorable  $^1\text{H}$ ,  $^{13}\text{C}$ , and  $^{15}\text{N}$  relaxation rates caused by the partial deuteration of the protein, constant-time triple-resonance NMR spectra can be acquired with higher digital resolution and sensitivity [Sattler, M. & Fesik, S. W. *Structure* **4:1245-1249 (1996)]**. In addition, various stable-isotopically labeled ( $^{15}\text{N}$  and  $^{13}\text{C}$  /  $^{15}\text{N}$ ) proteins can also be prepared using this procedure.

20

### Synthetic Polypeptides

- The term "polypeptide" is used in its broadest sense to refer to a compound of two or more subunit amino acids, amino acid analogs, or peptidomimetics. The subunits are
- 25 linked by peptide bonds. The terms "polypeptide", "protein", and "peptide" are used interchangeably herein, though preferably as used herein a "peptide" refers to a compound of at least two but less than fifty subunit amino acids, and a polypeptide or protein refers to compound of fifty or more amino acids. The polypeptides of the present invention may be chemically synthesized or as detailed above, genetically
- 30 engineered or isolated from natural sources.

In addition, potential drugs or agents that may be tested in the drug screening assays of the present invention may also be chemically synthesized. When the peptide is to be



modified, *e.g.*, acetylated, the modification can be at any time during the peptide synthesis, including using an acetyl-lysine as a starting material or acetylating a lysine residue of a peptide after the peptide has been synthesized. In the Example below, the acetyl-lysine-containing peptides were prepared on a MilliGen 9050 peptide  
 5 synthesizer (Perkin Elmer) using Fmoc/HBTU chemistry. Acetyl-lysine was incorporated using the reagent Fmoc-Ac-Lys with HBTU/DIPEA activation.

Thus, synthetic polypeptides, prepared using the well known techniques of solid phase, liquid phase, or peptide condensation techniques, or any combination thereof, can  
 10 include natural and unnatural amino acids. Amino acids used for peptide synthesis may be standard Boc ( $N^{\alpha}$ -amino protected  $N^{\alpha}$ -t-butyloxycarbonyl) amino acid resin with the standard deprotecting, neutralization, coupling and wash protocols of the original solid phase procedure of Merrifield [*J. Am. Chem. Soc.*, **85**:2149-2154 (1963)], or the base-labile  $N^{\alpha}$ -amino protected 9-fluorenylmethoxycarbonyl (Fmoc)  
 15 amino acids first described by Carpino and Han [*J. Org. Chem.*, **37**:3403-3409 (1972)]. Both Fmoc and Boc  $N^{\alpha}$ -amino protected amino acids can be obtained from Fluka, Bachem, Advanced Chemtech, Sigma, Cambridge Research Biochemical, Bachem, or Peninsula Labs or other chemical companies familiar to those who practice this art. In addition, the method of the invention can be used with other  $N^{\alpha}$ -protecting groups that  
 20 are familiar to those skilled in this art. Solid phase peptide synthesis may be accomplished by techniques familiar to those in the art and provided, for example, in Stewart and Young [Solid Phase Synthesis, Second Edition, Pierce Chemical Co., Rockford, IL (1984)] and Fields and Noble [*Int. J. Pept. Protein Res.*, **35**:161-214 (1990)], or using automated synthesizers, such as sold by ABS. Thus, polypeptides of  
 25 the invention may comprise D-amino acids, a combination of D- and L-amino acids, and various "designer" amino acids (*e.g.*,  $\beta$ -methyl amino acids,  $C\alpha$ -methyl amino acids, and  $N\alpha$ -methyl amino acids, etc.) to convey special properties. Synthetic amino acids include ornithine for lysine, fluorophenylalanine for phenylalanine, and norleucine for leucine or isoleucine. Additionally, by assigning specific amino acids at  
 30 specific coupling steps,  $\alpha$ -helices,  $\beta$  turns,  $\beta$  sheets,  $\gamma$ -turns, and cyclic peptides can be generated.

In a further embodiment, subunits of peptides that confer useful chemical and structural properties will be chosen. For example, peptides comprising D-amino acids will be resistant to L-amino acid-specific proteases *in vivo*. In addition, the present invention envisions preparing peptides that have more well defined structural

5 properties, and the use of peptidomimetics, and peptidomimetic bonds, such as ester bonds, to prepare peptides with novel properties. In another embodiment, a peptide may be generated that incorporates a reduced peptide bond, i.e.,  $R_1\text{-CH}_2\text{-NH-R}_2$ , where  $R_1$  and  $R_2$  are amino acid residues or sequences. A reduced peptide bond may be introduced as a dipeptide subunit. Such a molecule would be resistant to peptide bond

10 hydrolysis, *e.g.*, protease activity. Such peptides would provide ligands with unique function and activity, such as extended half-lives *in vivo* due to resistance to metabolic breakdown, or protease activity. Furthermore, it is well known that in certain systems constrained peptides show enhanced functional activity [Hruby, *Life Sciences*, **31**:189-199 (1982); Hruby *et al.*, *Biochem J.*, **268**:249-262 (1990)]; the present invention

15 provides a method to produce a constrained peptide that incorporates random sequences at all other positions.

*Constrained and cyclic peptides.* A constrained, cyclic or rigidized peptide may be prepared synthetically, provided that in at least two positions in the sequence of the

20 peptide an amino acid or amino acid analog is inserted that provides a chemical functional group capable of crosslinking to constrain, cyclise or rigidize the peptide after treatment to form the crosslink. Cyclization will be favored when a turn-inducing amino acid is incorporated. Examples of amino acids capable of crosslinking a peptide are cysteine to form disulfides, aspartic acid to form a lactone or a lactam, and a

25 chelator such as  $\gamma$ -carboxyl-glutamic acid (Gla) (Bachem) to chelate a transition metal and form a cross-link. Protected  $\gamma$ -carboxyl glutamic acid may be prepared by modifying the synthesis described by Zee-Cheng and Olson [*Biophys. Biochem. Res. Commun.*, **94**:1128-1132 (1980)]. A peptide in which the peptide sequence comprises at least two amino acids capable of crosslinking may be treated, *e.g.*, by oxidation of

30 cysteine residues to form a disulfide or addition of a metal ion to form a chelate, so as to crosslink the peptide and form a constrained, cyclic or rigidized peptide.

The present invention provides strategies to systematically prepare cross-links. For example, if four cysteine residues are incorporated in the peptide sequence, different protecting groups may be used (Hiskey, in *The Peptides: Analysis, Synthesis, Biology*, Vol. 3, Gross and Meienhofer, eds., Academic Press: New York, pp. 137-167 (1981); Ponsanti *et al.*, *Tetrahedron*, **46**:8255-8266 (1990)]. The first pair of cysteines may be deprotected and oxidized, then the second set may be deprotected and oxidized. In this way a defined set of disulfide cross-links may be formed. Alternatively, a pair of cysteines and a pair of chelating amino acid analogs may be incorporated so that the cross-links are of a different chemical nature.

10

*Non-classical amino acids that induce conformational constraints.* The following non-classical amino acids may be incorporated in the peptide in order to introduce particular conformational motifs: 1,2,3,4-tetrahydroisoquinoline-3-carboxylate [Kazmierski *et al.*, *J. Am. Chem. Soc.*, **113**:2275-2283 (1991)]; (2S,3S)-methyl-phenylalanine, (2S,3R)-methyl-phenylalanine, (2R,3S)-methyl-phenylalanine and (2R,3R)-methyl-phenylalanine (Kazmierski and Hruby, *Tetrahedron Lett.* (1991)]; 2-aminotetrahydronaphthalene-2-carboxylic acid [Landis, Ph.D. Thesis, University of Arizona (1989)]; hydroxy-1,2,3,4-tetrahydroisoquinoline-3-carboxylate [Miyake *et al.*, *J. Takeda Res. Labs.*, **43**:53-76 (1989)];  $\beta$ -carboline (D and L) [Kazmierski, Ph.D. Thesis, University of Arizona (1988)]; HIC (histidine isoquinoline carboxylic acid) [Zechel *et al.*, *Int. J. Pep. Protein Res.*, **43** (1991)]; and HIC (histidine cyclic urea) (Dharanipragada).

The following amino acid analogs and peptidomimetics may be incorporated into a peptide to induce or favor specific secondary structures: LL-Acp (LL-3-amino-2-propenidone-6-carboxylic acid), a  $\beta$ -turn inducing dipeptide analog [Kemp *et al.*, *J. Org. Chem.*, **50**:5834-5838 (1985)];  $\beta$ -sheet inducing analogs [Kemp *et al.*, *Tetrahedron Lett.*, **29**:5081-5082 (1988)];  $\beta$ -turn inducing analogs [Kemp *et al.*, *Tetrahedron Lett.*, **29**:5057-5060 (1988)];  $\alpha$ -helix inducing analogs (Kemp *et al.*, *Tetrahedron Lett.*, **29**:4935-4938 (1988)];  $\gamma$ -turn inducing analogs [Kemp *et al.*, *J. Org. Chem.*, **54**:109:115 (1989)]; and analogs provided by the following references: Nagai and Sato, *Tetrahedron Lett.*, **26**:647-650 (1985); DiMaio *et al.*, *J. Chem. Soc. Perkin Trans.*, p. 1687 (1989); also a Gly-Ala turn analog [Kahn *et al.*, *Tetrahedron*

*Lett.*, **30**:2317 (1989)]; amide bond isostere [Jones *et al.*, *Tetrahedron Lett.*, **29**:3853-3856 (1988)]; tretrazol [Zabrocki *et al.*, *J. Am. Chem. Soc.*, **110**:5875-5880 (1988)]; DTC [Samanen *et al.*, *Int. J. Protein Pep. Res.*, **35**:501:509 (1990)]; and analogs taught in Olson *et al.*, *J. Am. Chem. Sci.*, **112**:323-333 (1990) and Garvey *et al.*, *J. Org. Chem.*, **56**:436 (1990). Conformationally restricted mimetics of beta turns and beta bulges, and peptides containing them, are described in U.S. Patent No. 5,440,013, issued August 8, 1995 to Kahn.

### Structure-based Mutation Analysis

10

Protein structural analysis using NMR spectroscopy has several unique advantages. In addition to high-resolution three-dimensional structural information, the chemical shift assignments for the protein obtained in the structural study further provides a map of the entire protein at the atomic level, which can be used for structure-based

15

biochemical analysis of protein-protein interactions. For example, the information generated from the NMR structural analysis can also serve to identify specific amino acid residues in the peptide-binding site for complementary mutagenesis studies.

Specific focus can be placed on those residues that display long-range NOEs (particularly the side-chain NOEs in the  $^{13}\text{C}$ -NOESY data) between the bromodomain

20

and a peptide comprising an acetyl-lysine.

To ensure mutant proteins are valid for functional analysis, it can be determined as to whether a mutation results in any significant perturbation of the overall conformation of the bromodomain, particularly the effects of mutation on the acetyl-lysine binding sites. NMR spectroscopy is a powerful method for examining the effects of such a mutation on the conformation of the protein. One can readily obtain information about the global conformation of a mutant protein from the proton ( $^1\text{H}$ ) 1D spectrum, by examining the chemical shift dispersion and peak line-width of NMR signals of amide, aromatic and aliphatic protons. Moreover, 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra reveal details of the effects of a mutation on both local and global conformation of the protein, since every single  $^1\text{H}/^{15}\text{N}$  signal (both the chemical shift and line-shape) in the NMR spectrum is a "reporter" for a particular amino acid residue. Thus, to assess how mutations effect protein stability and the overall protein conformation, the  $^{15}\text{N}$  HSQC

spectra of mutated proteins can be compared to that of the wild-type protein bromodomain.

Chemical-shift perturbations due to ligand binding have proven to be a reliable and sensitive probe for the ligand binding site of the protein. This is because the chemical-shift changes of the backbone amide groups are likely to reflect any changes in protein conformation and/or hydrogen bonding due to the peptide/ligand binding. To examine the effects of a mutation on the ligand binding (in this case the ligand is a peptide comprising an acetyl-lysine), peptide titration experiments can be conducted by following the changes of  $^1\text{H}/^{15}\text{N}$  signals of the mutant proteins as a function of the peptide concentration. These experiments indicate whether the acetyl-lysine binding site remains the same or changes in the mutants relative to the wild type protein. The effects of the mutation on the peptide binding affinity can also be examined by NMR spectroscopy. If the mutated proteins result in the reduction of the binding affinity, a change of the exchange phenomenon between the free and the ligand-bound signals should be observed in NMR spectrum. If the reduction in binding affinity causes the peptide binding to change from a slow exchange rate to a fast exchange rate, on the NMR time scale, then the peptide binding affinity can be determined from the NMR titration experiment. From these mutation analyses key amino acid residues that are important for binding a peptide comprising the acetyl-lysine can be identified. Such analysis has been exemplified below.

#### Protein Structure Determination by NMR Spectroscopy

The NMR results from the present invention are summarized by the atomic structure coordinates of the free form of the P/CAF bromodomain (Table 5) and of the P/CAF bromodomain-acetyl-histamine complex (Table 6). The NMR chemical shift assignments of the P/CAF bromodomain are included in the chemical shift table (Table 1) for the  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum of P/CAF bromodomain. The unambiguous NOE-derived Inter-proton Distance Restraints are in Table 2, the ambiguous NOE-derived Inter-proton Distance Restraints are in Table 3, and the  $^1\text{H}$  bonding restraints are disclosed in Table 4.

5 (1994)]. The water flip-back scheme is used in these NMR pulse programs to minimize amide signal attenuation from water exchange. Sequential side-chain assignments are typically accomplished from a series of 3D NMR experiments with alternative approaches to confirm the assignments. These experiments include 3D  $^{15}\text{N}$  TOCSY-HSQC, HCCH-TOCSY, (H)C(CO)NH-TOCSY, and H(C)(CO)NH-TOCSY  
10 [see Clore, G. M. & Gronenborn, A. M. *Meth. Enzymol.* **239**:249-363 (1994); Sattler *et al.*, *Prog. in Nuclear Magnetic Resonance Spec.* **4**:93-158 (1999)].

15 splitting using a fractionally  $^{13}\text{C}$ -labeled protein sample, which can be readily prepared using M9 minimal medium containing 10%  $^{13}\text{C}$ -/90%  $^{12}\text{C}$ -glucose mixture [see Neri, *et al.*, *Biochemistry* **28**:7510-7516 (1989)].

25  $\beta$  methylene protons.

30

**NOE Distance Restraints:** Distance restraints are obtained from analysis of  $^{15}\text{N}$ , and  $^{13}\text{C}$ -edited 3D NOESY data, which can be collected with different mixing times to minimize spin diffusion problems. The nuclear Overhauser effect (NOE)-derived

restraints are categorized as strong (1.8-3 Å), medium (1.8-4 Å) or weak (1.8-5 Å) based on the observed NOE intensities. A recently developed procedure for the iterative automated NOE analysis by using ARIA [see Nilges *et al.*, *Prog. NMR Spectroscopy* **32**:107-139 (1998)] can be employed which integrates with X-PLOR for structural calculations. To ensure the success of ARIA/X-PLOR-assisted NOE analysis and structure calculations, the ARIA assigned NOE peaks can be manually confirmed.

*Intermolecular NOE Distance Restraints:* For the structural determination of a protein/peptide complex, intermolecular NOE distance restraints can be obtained from a  $^{13}\text{C}$ -edited ( $F_1$ ) and  $^{15}\text{N}$ , and  $^{13}\text{C}$ -filtered ( $F_3$ ) 3D NOESY data set collected for a sample containing isotope-labeled protein and non-labeled peptide.

*Structure Calculations and Refinements:* Structures of the protein can be generated using a distance geometry/simulated annealing protocol with the X-PLOR program [see Nilges, *et al.*, *FEBS Lett.* **229**:317-324 (1988); Kuszewski, *et al.*, *J. Biomol. NMR* **2**:33-56 (1992); Brünger, A. T. *X-PLOR Version 3.1: A system for X-Ray crystallography and NMR* (Yale University Press, New Haven, CT, 1993)]. The structure calculations can employ inter-proton distance restraints obtained from  $^{15}\text{N}$ - and  $^{13}\text{C}$ -resolved NOESY spectra. The initial low-resolution structures can be used to facilitate NOE assignments, and help identify hydrogen bonding partners for slowly exchanging amide protons. The experimental restraints of dihedral angles and hydrogen bonds can be included in the distance restraints for structure refinements.

#### Protein-Structure Based Design of Agonists and Antagonists of the Bromodomain-Acetyl-Lysine Binding Complex

Once the three-dimensional structure of the Bromodomain and the Bromodomain-acetyl-lysine binding complex are determined, a potential drug or agent (antagonist or agonist) can be examined through the use of computer modeling using a docking program such as GRAM, DOCK, or AUTODOCK [Dunbrack *et al.*, 1997, *supra*]. This procedure can include computer fitting of potential agents to the bromodomain, for example, to ascertain how well the shape and the chemical structure of the potential ligand will complement or interfere with the interaction between the bromodomain and

5

15

20

25

30



5

10

15

25

30

Phage libraries have been constructed which when infected into host *E. coli* produce random peptide sequences of approximately 10 to 15 amino acids [Parmley and Smith,

Gene 73:305-318 (1988), Scott and Smith, Science 249:386-249 (1990)]. Specifically, the phage library can be mixed in low dilutions with permissive *E. coli* in low melting point LB agar which is then poured on top of LB agar plates. After incubating the plates at 37°C for a period of time, small clear plaques in a lawn of *E. coli* will form which represents active phage growth and lysis of the *E. coli*. A representative of these phages can be absorbed to nylon filters by placing dry filters onto the agar plates. The filters can be marked for orientation, removed, and placed in washing solutions to block any remaining absorbent sites. The filters can then be placed in a solution containing, for example, a radioactive bromodomain. After a specified incubation period, the filters can be thoroughly washed and developed for autoradiography. Plaques containing the phage that bind to the radioactive bromodomain can then be identified. These phages can be further cloned and then retested for their ability to bind to the bromodomain as before. Once the phage has been purified, the binding sequence contained within the phage can be determined by standard DNA sequencing techniques. Once the DNA sequence is known, synthetic peptides can be generated which are encoded by these sequences. These peptides can be tested, for example, for their ability to modulate the affinity of the bromodomain for its binding partner (*e.g.*, a protein comprising an acetyl-lysine or a fragment of that protein).

The effective peptide(s) can be synthesized in large quantities for use in *in vivo* models and eventually in humans to treat certain tumors. It should be emphasized that synthetic peptide production is relatively non-labor intensive, easily manufactured, quality controlled and thus, large quantities of the desired product can be produced quite cheaply. Similar combinations of mass produced synthetic peptides have been used with great success [Patarroyo, *Vaccine*, 10:175-178 (1990)].

### Drug Screening Assays

The drug screening assays of the present invention may use any of a number of means for determining the interaction between an agent or drug and a peptide comprising an acetyl-lysine and/or a bromodomain. Thus, standard high throughput drug screening procedures can be employed using a library of low molecular weight compounds, for

example that can be screened to identify a binding partner for the bromodomain. Any such chemical library can be used including those discussed above.

In a particular assay, a bromodomain is placed on or coated onto a solid support.

- 5 Methods for placing the peptides or proteins on the solid support are well known in the art and include such things as linking biotin to the protein and linking avidin to the solid support. An agent is allowed to equilibrate with the bromodomain to test for binding. Generally, the solid support is washed and agents that are retained are selected as potential drugs. Alternatively, a peptide comprising an acetyl-lysine is  
10 placed on or coated onto a solid support. In a particular embodiment of this type, the peptide comprises the amino acid sequence of SEQ ID NO:4.

- The agent may be labeled. For example, in one embodiment radiolabeled agents are used to measure the binding of the agent. In another embodiment the agents have  
15 fluorescent markers. In yet another embodiment, a Biocore chip (Pharmacia) coated with the bromodomain is used, for example and the change in surface conductivity can be measured.

- In addition, since a number of proteins have been identified that contain  
20 bromodomains, and the binding partners of many of these proteins are known, the fact that the bromodomain specifically binds to an acetylated lysine as disclosed herein allows the identification and preparation of a number of potential modulators of the bromodomain-acetyl-lysine binding complex based on the amino acid sequences of the binding partners to the proteins. Such potential modulators include : ISYGR-AcK-  
25 KRRQRR (SEQ ID NO:4), ARKSTGG-AcK-APRKQL (SEQ ID NO:5) and QSTSRHK-AcK-LMFKTE (SEQ ID NO:6) which bind to the P/CAF bromodomain as shown in the Example, below. Such peptides also can be used, for example, as a starting point for the design of an inhibitor of the bromodomain-acetyl-lysine binding complex.

30

Alternatively, a drug can be specifically designed to bind to the ZA loop of a bromodomain for example, such as the P/CAF bromodomain, and be assayed through NMR based methodology [Shuker *et al.*, *Science* 274:1531-1534 (1996) hereby

incorporated by reference in its entirety.] In a particular embodiment, analogs of the binding partner of the bromodomain can be used in this analysis. One such peptide has the amino acid sequence of SEQ ID NO:4. In another embodiment of this type, the peptide has the amino acid sequence of SEQ ID NO:5. In another such embodiment of  
 5 this type, the peptide has the amino acid sequence of SEQ ID NO:6.

The assay begins with contacting a compound with a  $^{15}\text{N}$ -labeled bromodomain. Binding of the compound with the ZA loop of the bromodomain can be determined by monitoring the  $^{15}\text{N}$ - or  $^1\text{H}$ -amide chemical shift changes in two dimensional  $^{15}\text{N}$ -  
 10 heteronuclear single-quantum correlation ( $^{15}\text{N}$ -HSQC) spectra upon the addition of the compound to the  $^{15}\text{N}$ -labeled bromodomain. Since these spectra can be rapidly obtained, it is feasible to screen a large number of compounds [Shuker *et al.*, *Science* **274**:1531-1534 (1996)]. A compound is identified as a potential ligand if it binds to the ZA loop of the bromodomain. In a further embodiment, the potential ligand can  
 15 then be used as a model structure, and analogs to the compound can be obtained (e.g, from the vast chemical libraries commercially available, or alternatively through *de novo* synthesis). The analogs are then screened for their ability to bind the ZA loop of the bromodomain thus to obtain a ligand. An analog of the potential ligand is chosen as a ligand when it binds to the ZA loop of the bromodomain with a higher binding  
 20 affinity than the potential ligand. In a preferred embodiment of this type the analogs are screened by monitoring the  $^{15}\text{N}$ - or  $^1\text{H}$ -amide chemical shift changes in two dimensional  $^{15}\text{N}$ -heteronuclear single-quantum correlation ( $^{15}\text{N}$ -HSQC) spectra upon the addition of the analog to the  $^{15}\text{N}$ -labeled bromodomain as described above.

25 In another further embodiment, compounds are screened for binding to two nearby sites on the bromodomain. In this case, a compound that binds a first site of the bromodomain does not bind a second nearby site. Binding to the second site can be determined by monitoring changes in a different set of amide chemical shifts in either the original screen or a second screen conducted in the presence of a ligand (or  
 30 potential ligand) for the first site. From an analysis of the chemical shift changes the approximate location of a potential ligand for the second site is identified. Optimization of the second ligand for binding to the site is then carried out by screening structurally related compounds (e.g., analogs as described above). When

- ligands for the first site and the second site are identified, their location and orientation in the ternary complex can be determined experimentally either by NMR spectroscopy or X-ray crystallography. On the basis of this structural information, a linked compound is synthesized in which the ligand for the first site and the ligand for the
- 5 second site are linked. In a preferred embodiment of this type the two ligands are covalently linked. This linked compound is tested to determine if it has a higher binding affinity for the bromodomain than either of the two individual ligands. A linked compound is selected as a ligand when it has a higher binding affinity for the bromodomain than either of the two ligands. In a preferred embodiment the affinity of
- 10 the linked compound with the bromodomain is determined monitoring the  $^{15}\text{N}$ - or  $^1\text{H}$ -amide chemical shift changes in two dimensional  $^{15}\text{N}$ -heteronuclear single-quantum correlation ( $^{15}\text{N}$ -HSQC) spectra upon the addition of the linked compound to the  $^{15}\text{N}$ -labeled bromodomain as described above.
- 15 A larger linked compound can be constructed in an analogous manner, *e.g.*, linking three ligands which bind to three nearby sites on the bromodomain to form a multilinked compound that has an even higher affinity for the bromodomain than the linked compound.

## 20 Identification of New Bromodomains

- By disclosing that protein bound acetyl-lysine is a binding partner for bromodomains, the present invention provides a method of identifying novel proteins that contain bromodomains. In short, a protein fragment or analog thereof comprising an acetyl-
- 25 lysine can be used as bait to identify a binding partner that comprises a bromodomain. Any one of a number of procedures can be carried out to identify such a binding partner. One such assay comprises passing a cell extract over the bait peptide which is attached to a solid support. After washing the solid support to remove any non-specific binders, the bromodomain containing protein can be eluted from the solid
- 30 support with an appropriate eluant. In a particular embodiment, the free bait peptide can be used in the elution. Other methodology includes the use of a yeast two-hybrid system, a GST pull down assay, ELISA, immunometric assays, and a modification of the CORT procedure of Schlessinger *et al.*, (US Patent No. 5,858,686, Issued on

January 12, 1999 which is hereby incorporated by reference in its entirety) for use with the bromodomain-acetyl-lysine binding complex.

Labels:

5

Suitable labels include enzymes, fluorophores (*e.g.*, fluorescein isothiocyanate (FITC), phycoerythrin (PE), Texas red (TR), rhodamine, free or chelated lanthanide series salts, especially  $\text{Eu}^{3+}$ , to name a few fluorophores), chromophores, radioisotopes, chelating agents, dyes, colloidal gold, latex particles, ligands (*e.g.*, biotin), and

10 chemiluminescent agents. When a control marker is employed, the same or different labels may be used for the test and control marker gene.

In the instance where a radioactive label, such as the isotopes  $^3\text{H}$ ,  $^{14}\text{C}$ ,  $^{32}\text{P}$ ,  $^{35}\text{S}$ ,  $^{36}\text{Cl}$ ,  $^{51}\text{Cr}$ ,  $^{57}\text{Co}$ ,  $^{58}\text{Co}$ ,  $^{59}\text{Fe}$ ,  $^{90}\text{Y}$ ,  $^{125}\text{I}$ ,  $^{131}\text{I}$ , and  $^{186}\text{Re}$  are used, known currently available

15 counting procedures may be utilized. In the instance where the label is an enzyme, detection may be accomplished by any of the presently utilized colorimetric, spectrophotometric, fluorospectrophotometric, amperometric or gasometric techniques known in the art.

20 Direct labels are one example of labels which can be used according to the present invention. A direct label has been defined as an entity, which in its natural state, is readily visible, either to the naked eye, or with the aid of an optical filter and/or applied stimulation, *e.g.* U.V. light to promote fluorescence. Among examples of colored labels, which can be used according to the present invention, include metallic sol

25 particles, for example, gold sol particles such as those described by Leuvering (U.S. Patent 4,313,734); dye sol particles such as described by Gribnau *et al.* (U.S. Patent 4,373,932 and May *et al.* (WO 88/08534); dyed latex such as described by May, *supra*, Snyder (EP-A 0 280 559 and 0 281 327); or dyes encapsulated in liposomes as described by Campbell *et al.* (U.S. Patent 4,703,017). Other direct labels include a

30 radionucleotide, a fluorescent moiety or a luminescent moiety. In addition to these direct labeling devices, indirect labels comprising enzymes can also be used according to the present invention. Various types of enzyme linked immunoassays are well known in the art, for example, alkaline phosphatase and horseradish peroxidase,

002220" 47E07560

lysozyme, glucose-6-phosphate dehydrogenase, lactate dehydrogenase, urease, these and others have been discussed in detail by Eva Engvall in Enzyme Immunoassay ELISA and EMIT in *Methods in Enzymology*, **70**:419-439 (1980) and in U.S. Patent 4,857,453.

5

Suitable enzymes include, but are not limited to, alkaline phosphatase,  $\beta$ -galactosidase, green fluorescent protein and its derivatives, luciferase, and horseradish peroxidase.

Other labels for use in the invention include magnetic beads or magnetic resonance  
10 imaging labels.

#### Antibodies to Portions of the Bromodomain that Interact with Acetyl-Lysine

According to the present invention, the bromodomains, and more particularly the ZA  
15 loops of the bromodomains and fragments thereof can be produced by a recombinant source, or through chemical synthesis, or through the modification of these peptides and fragments; and derivatives or analogs thereof, including fusion proteins, may be used as an immunogen to generate antibodies that specifically interfere with the formation of the bromodomain-acetyl-lysine binding complex. Similarly, antibodies  
20 can be raised against peptides that comprise one or more acetyl-lysine residues which also interfere with the formation of the bromodomain-acetyl-lysine binding complex. Such antibodies include but are not limited to polyclonal, monoclonal, chimeric, single chain, Fab fragments, and a Fab expression library.

25 Various procedures known in the art may be used for the production of the polyclonal antibodies. For the production of antibody, various host animals can be immunized by injection with the peptide having the amino acid sequence of SEQ ID NO:3, for example, or a derivative (*e.g.*, or fusion protein) thereof, including but not limited to rabbits, mice, rats, sheep, goats, etc. In one embodiment, the peptide can be  
30 conjugated to an immunogenic carrier, *e.g.*, bovine serum albumin (BSA) or keyhole limpet hemocyanin (KLH). Various adjuvants may be used to increase the immunological response, depending on the host species, including but not limited to Freund's (complete and incomplete), mineral gels such as aluminum hydroxide, surface

active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanins, dinitrophenol, and potentially useful human adjuvants such as BCG (*bacille Calmette-Guerin*) and *Corynebacterium parvum*.

- 5 For preparation of monoclonal antibodies directed toward the peptides or protein fragments of the present invention, or analog, or derivative thereof, any technique that provides for the production of antibody molecules by continuous cell lines in culture may be used. These include but are not limited to the hybridoma technique originally developed by Kohler and Milstein [*Nature*, **256**:495-497 (1975)], as well as the trioma
- 10 technique, the human B-cell hybridoma technique [Kozbor *et al.*, *Immunology Today*, **4**:72 (1983); Cote *et al.*, *Proc. Natl. Acad. Sci. U.S.A.*, **80**:2026-2030 (1983)], and the EBV-hybridoma technique to produce human monoclonal antibodies [Cole *et al.*, in *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc., pp. 77-96 (1985)]. In an additional embodiment of the invention, monoclonal antibodies can be produced in
- 15 germ-free animals utilizing technology described in PCT/US90/02545. In fact, according to the invention, techniques developed for the production of "chimeric antibodies" [Morrison *et al.*, *J. Bacteriol.*, **159**:870 (1984); Neuberger *et al.*, *Nature*, **312**:604-608 (1984); Takeda *et al.*, *Nature*, **314**:452-454 (1985)] by splicing the genes from a mouse antibody molecule specific for the peptide having the amino acid
- 20 sequence of SEQ ID NO:3, for example, together with genes from a human antibody molecule of appropriate biological activity can be used; such antibodies are within the scope of this invention. Such human or humanized chimeric antibodies are preferred for use in therapy of human diseases or disorders (described *infra*), since the human or humanized antibodies are much less likely than xenogenic antibodies to induce an
- 25 immune response, in particular an allergic response, themselves.

According to the invention, techniques described for the production of single chain antibodies [U.S. Patent Nos. 5,476,786 and 5,132,405 to Huston; U.S. Patent 4,946,778] can be adapted to produce specific single chain antibodies. An additional

30 embodiment of the invention utilizes the techniques described for the construction of Fab expression libraries [Huse *et al.*, *Science*, **246**:1275-1281 (1989)] to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity.



Antibody fragments which contain the idiotype of the antibody molecule can be generated by known techniques. For example, such fragments include but are not limited to: the  $F(ab')_2$  fragment which can be produced by pepsin digestion of the antibody molecule; the Fab' fragments which can be generated by reducing the disulfide bridges of the  $F(ab')_2$  fragment, and the Fab fragments which can be generated by treating the antibody molecule with papain and a reducing agent.

In the production of antibodies, screening for the desired antibody can be accomplished by techniques known in the art, *e.g.*, radioimmunoassay, ELISA (enzyme-linked immunosorbant assay), "sandwich" immunoassays, immunoradiometric assays, gel diffusion precipitin reactions, immunodiffusion assays, *in situ* immunoassays (using colloidal gold, enzyme or radioisotope labels, for example), western blots, precipitation reactions, agglutination assays (*e.g.*, gel agglutination assays, hemagglutination assays), complement fixation assays, immunofluorescence assays, protein A assays, and immunoelectrophoresis assays, etc. In one embodiment, antibody binding is detected by detecting a label on the primary antibody. In another embodiment, the primary antibody is detected by detecting binding of a secondary antibody or reagent to the primary antibody. In a further embodiment, the secondary antibody is labeled. Many means are known in the art for detecting binding in an immunoassay and are within the scope of the present invention. For example, to select antibodies which recognize a specific epitope of a ZA loop of a bromodomain, for example, one may assay generated hybridomas for a product which binds to a bromodomain fragment containing such an epitope and choose those which do not cross-react with bromodomain fragments that do not include that epitope.

25

In a specific embodiment, antibodies that interfere with the formation of the bromodomain-acetyl-lysine complex can be generated. Such antibodies can be tested using the assays described and could potentially be used in anti-cancer therapies.

30

#### Administration

According to the invention, the component or components of a therapeutic composition, *e.g.*, an agent of the invention that interferes with the bromodomain-

acetyl-lysine binding complex such as the peptide having the amino acid sequence of SEQ ID NOs:4, 5, or 6 and a pharmaceutically acceptable carrier, may be introduced parenterally, transmucosally, *e.g.*, orally, nasally, or rectally, or transdermally.

Preferably, administration is parenteral, *e.g.*, via intravenous injection, and also

- 5 including, but is not limited to, intra-arteriole, intramuscular, intradermal, subcutaneous, intraperitoneal, intraventricular, and intracranial administration.

In a preferred aspect, the agent of the present invention can cross cellular and nuclear membranes, which would allow for intravenous or oral administration. Strategies are  
10 available for such crossing, including but not limited to, increasing the hydrophobic nature of a molecule; introducing the molecule as a conjugate to a carrier, such as a ligand to a specific receptor, targeted to a receptor; and the like.

The present invention also provides for conjugating targeting molecules to such an

- 15 agent. "Targeting molecule" as used herein shall mean a molecule which, when administered *in vivo*, localizes to desired location(s). In various embodiments, the targeting molecule can be a peptide or protein, antibody, lectin, carbohydrate, or steroid. In one embodiment, the targeting molecule is a peptide ligand of a receptor on the target cell. In a specific embodiment, the targeting molecule is an antibody.  
20 Preferably, the targeting molecule is a monoclonal antibody. In one embodiment, to facilitate crosslinking the antibody can be reduced to two heavy and light chain heterodimers, or the F(ab')<sub>2</sub> fragment can be reduced, and crosslinked to the agent via the reduced sulfhydryl. Antibodies for use as targeting molecule are specific for a cell surface antigen.

25

In another embodiment, the therapeutic compound can be delivered in a vesicle, in particular a liposome [*see* Langer, *Science*, **249**:1527-1533 (1990); Treat *et al.*, in *Liposomes in the Therapy of Infectious Disease and Cancer*, Lopez-Berestein and Fidler (eds.), Liss: New York, pp. 353-365 (1989); Lopez-Berestein, *ibid.*, pp. 317-  
30 327; see generally *ibid.*].

In yet another embodiment, the therapeutic compound can be delivered in a controlled release system. For example, the agent may be administered using intravenous

002220" 42507560

infusion, an implantable osmotic pump, a transdermal patch, liposomes, or other modes of administration. In one embodiment, a pump may be used [see Langer, *supra*; Sefton, *CRC Crit. Ref. Biomed. Eng.*, **14**:201 (1987); Buchwald *et al.*, *Surgery*, **88**:507 (1980); Saudek *et al.*, *N. Engl. J. Med.*, **321**:574 (1989)]. In another embodiment,

5 polymeric materials can be used [see *Medical Applications of Controlled Release*, Langer and Wise (eds.), CRC Press: Boca Raton, Florida (1974); *Controlled Drug Bioavailability, Drug Product Design and Performance*, Smolen and Ball (eds.), Wiley: New York (1984); Ranger and Peppas, *J. Macromol. Sci. Rev. Macromol. Chem.*, **23**:61 (1983); see also Levy *et al.*, *Science*, **228**:190 (1985); During *et al.*, *Ann.*

10 *Neurol.*, **25**:351 (1989); Howard *et al.*, *J. Neurosurg.*, **71**:105 (1989)]. In yet another embodiment, a controlled release system can be placed in proximity of the therapeutic target, *i.e.*, the bone marrow, thus requiring only a fraction of the systemic dose [see, *e.g.*, Goodson, in *Medical Applications of Controlled Release, supra*, vol. 2, pp. 115-138 (1984)]. Other controlled release systems are discussed in the review by Langer

15 [*Science*, **249**:1527-1533 (1990)].

*Pharmaceutical Compositions.* In yet another aspect of the present invention, provided are pharmaceutical compositions of the above. Such pharmaceutical compositions may be for administration for injection, or for oral, pulmonary, nasal or other forms of

20 administration. In general, comprehended by the invention are pharmaceutical compositions comprising effective amounts of a low molecular weight component or components, or derivative products, of the invention together with pharmaceutically acceptable diluents, preservatives, solubilizers, emulsifiers, adjuvants and/or carriers. Such compositions include diluents of various buffer content (*e.g.*, Tris-HCl, acetate,

25 phosphate), pH and ionic strength; additives such as detergents and solubilizing agents (*e.g.*, Tween 80, Polysorbate 80), anti-oxidants (*e.g.*, ascorbic acid, sodium metabisulfite), preservatives (*e.g.*, Thimersol, benzyl alcohol) and bulking substances (*e.g.*, lactose, mannitol); incorporation of the material into particulate preparations of polymeric compounds such as polylactic acid, polyglycolic acid, etc. or into liposomes.

30 Hylauronic acid may also be used. Such compositions may influence the physical state, stability, rate of *in vivo* release, and rate of *in vivo* clearance of the present proteins and derivatives. See, *e.g.*, Remington's Pharmaceutical Sciences, 18th Ed. [1990, Mack Publishing Co., Easton, PA 18042] pages 1435-1712 which are herein

09510314 022200

incorporated by reference. The compositions may be prepared in liquid form, or may be in dried powder, such as lyophilized form.

*Oral Delivery.* Contemplated for use herein are oral solid dosage forms, which are described generally in Remington's Pharmaceutical Sciences, 18th Ed. 1990 (Mack Publishing Co. Easton PA 18042) at Chapter 89, which is herein incorporated by reference. Solid dosage forms include tablets, capsules, pills, troches or lozenges, cachets or pellets. Also, liposomal or proteinoid encapsulation may be used to formulate the present compositions (as, for example, proteinoid microspheres reported in U.S. Patent No. 4,925,673). Liposomal encapsulation may be used and the liposomes may be derivatized with various polymers (e.g., U.S. Patent No. 5,013,556). A description of possible solid dosage forms for the therapeutic is given by Marshall, K. In: *Modern Pharmaceutics* Edited by G.S. Banker and C.T. Rhodes Chapter 10, 1979, herein incorporated by reference. In general, the formulation will include an agent of the present invention (or chemically modified forms thereof) and inert ingredients which allow for protection against the stomach environment, and release of the biologically active material in the intestine.

Also specifically contemplated are oral dosage forms of the above derivatized component or components. The component or components may be chemically modified so that oral delivery of the derivative is efficacious. Generally, the chemical modification contemplated is the attachment of at least one moiety to the component molecule itself, where said moiety permits (a) inhibition of proteolysis; and (b) uptake into the blood stream from the stomach or intestine. Also desired is the increase in overall stability of the component or components and increase in circulation time in the body. An example of such a moiety is polyethylene glycol.

For the component (or derivative) the location of release may be the stomach, the small intestine (the duodenum, the jejunum, or the ileum), or the large intestine. One skilled in the art has available formulations which will not dissolve in the stomach, yet will release the material in the duodenum or elsewhere in the intestine. Preferably, the release will avoid the deleterious effects of the stomach environment, either by

protection of the protein (or derivative) or by release of the biologically active material beyond the stomach environment, such as in the intestine.

The therapeutic can be included in the formulation as fine multi-particulates in the form of granules or pellets of particle size about 1 mm. The formulation of the material for capsule administration could also be as a powder, lightly compressed plugs or even as tablets. The therapeutic could be prepared by compression.

One may dilute or increase the volume of the therapeutic with an inert material. These diluents could include carbohydrates, especially mannitol,  $\alpha$ -lactose, anhydrous lactose, cellulose, sucrose, modified dextrans and starch. Certain inorganic salts may be also be used as fillers including calcium triphosphate, magnesium carbonate and sodium chloride. Some commercially available diluents are Fast-Flo, Emdex, STA-Rx 1500, Emcompress and Avicell.

Disintegrants may be included in the formulation of the therapeutic into a solid dosage form. Materials used as disintegrates include but are not limited to starch, including the commercial disintegrant based on starch, Explotab. Binders also may be used to hold the therapeutic agent together to form a hard tablet and include materials from natural products such as acacia, tragacanth, starch and gelatin.

An anti-frictional agent may be included in the formulation of the therapeutic to prevent sticking during the formulation process. Lubricants may be used as a layer between the therapeutic and the die wall. Glidants that might improve the flow properties of the drug during formulation and to aid rearrangement during compression also might be added. The glidants may include starch, talc, pyrogenic silica and hydrated silicoaluminate.

In addition, to aid dissolution of the therapeutic into the aqueous environment a surfactant might be added as a wetting agent. Additives which potentially enhance uptake of the protein (or derivative) are for instance the fatty acids oleic acid, linoleic acid and linolenic acid.

*Nasal Delivery.* Nasal delivery of an agent of the present invention (or derivative) is also contemplated. Nasal delivery allows the passage of a peptide, for example, to the blood stream directly after administering the therapeutic product to the nose, without the necessity for deposition of the product in the lung. Formulations for nasal delivery  
5 include those with dextran or cyclodextran.

*Transdermal administration.* Various and numerous methods are known in the art for transdermal administration of a drug, *e.g.*, via a transdermal patch. Transdermal patches are described in for example, U.S. Patent No. 5,407,713, issued April 18, 1995  
10 to Rolando *et al.*; U.S. Patent No. 5,352,456, issued October 4, 1994 to Fallon *et al.*; U.S. Patent No. 5,332,213 issued August 9, 1994 to D'Angelo *et al.*; U.S. Patent No. 5,336,168, issued August 9, 1994 to Sibalis; U.S. Patent No. 5,290,561, issued March 1, 1994 to Farhadieh *et al.*; U.S. Patent No. 5,254,346, issued October 19, 1993 to Tucker *et al.*; U.S. Patent No. 5,164,189, issued November 17, 1992 to Berger *et al.*;  
15 U.S. Patent No. 5,163,899, issued November 17, 1992 to Sibalis; U.S. Patent Nos. 5,088,977 and 5,087,240, both issued February 18, 1992 to Sibalis; U.S. Patent No. 5,008,110, issued April 16, 1991 to Benecke *et al.*; and U.S. Patent No. 4,921,475, issued May 1, 1990 to Sibalis, the disclosure of each of which is incorporated herein by reference in its entirety.

20 It can be readily appreciated that a transdermal route of administration may be enhanced by use of a dermal penetration enhancer, *e.g.*, such as enhancers described in U.S. Patent No. 5,164,189 (*supra*), U.S. Patent No. 5,008,110 (*supra*), and U.S. Patent No. 4,879,119, issued November 7, 1989 to Aruga *et al.*, the disclosure of each of  
25 which is incorporated herein by reference in its entirety.

*Pulmonary Delivery.* Also contemplated herein is pulmonary delivery of the pharmaceutical compositions of the present invention. A pharmaceutical composition of the present invention is delivered to the lungs of a mammal while inhaling and  
30 traverses across the lung epithelial lining to the blood stream. Other reports of this include Adjei *et al.* [*Pharmaceutical Research*, 7:565-569 (1990); Adjei *et al.*, *International Journal of Pharmaceutics*, 63:135-144 (1990) (leuprolide acetate); Braquet *et al.*, *Journal of Cardiovascular Pharmacology*, 13(suppl. 5):143-146 (1989)

(endothelin-1); Hubbard *et al.*, *Annals of Internal Medicine*, **Vol. III**, pp. 206-212 (1989) ( $\alpha$ 1-antitrypsin); Smith *et al.*, *J. Clin. Invest.*, **84**:1145-1146 (1989) ( $\alpha$ -1-proteinase); Oswein *et al.*, "Aerosolization of Proteins", *Proceedings of Symposium on Respiratory Drug Delivery II*, Keystone, Colorado, March, (1990) (recombinant human growth hormone); Debs *et al.*, *J. Immunol.*, **140**:3482-3488 (1988) (interferon- $\gamma$  and tumor necrosis factor alpha); Platz *et al.*, U.S. Patent No. 5,284,656 (granulocyte colony stimulating factor)]. A method and composition for pulmonary delivery of drugs for systemic effect is described in U.S. Patent No. 5,451,569, issued September 19, 1995 to Wong *et al.*

10

A subject in whom administration of an agent of the present invention is an effective therapeutic regiment for cancer, for example, is preferably a human, but can be any animal. Thus, as can be readily appreciated by one of ordinary skill in the art, the methods and pharmaceutical compositions of the present invention are particularly suited to administration to any animal, *e.g.*, for veterinary medical use, particularly for a mammal, and including, but by no means limited to, domestic animals, such as feline or canine subjects, farm animals, including bovine, equine, caprine, ovine, and porcine subjects, wild animals (whether in the wild or in a zoological garden), research animals, such as mice, rats, rabbits, goats, sheep, pigs, dogs, cats, avian species, such as chickens, turkeys, and songbirds.

15  
20

The present invention may be better understood by reference to the following non-limiting Example, which is provided as exemplary of the invention. The following example is presented in order to more fully illustrate the preferred embodiments of the invention. It should in no way be construed, however, as limiting the broad scope of the invention.

25

30

EXAMPLE  
STRUCTURE AND LIGAND OF A HISTONE  
ACETYLTRANSFERASE BROMODOMAIN

5 Introduction

The bromodomain is a protein motif comprising approximately 110 amino acids that is found in practically all nuclear histone acetyltransferases (HATs) [Jeanmougin *et al.*, Trends in Biochemical Sciences, **22**:151-153 (1997)]. However, despite the seemingly requisite occurrence of this motif in HATs, their role in these enzymes is unknown.

10 Indeed, although this motif has also been identified in other chromatin proteins, heretofore not even one binding partner for a bromodomain had been identified.

Materials and Methods

*Sample preparation:* The bromodomain of P/CAF (residues 719-832 of SEQ ID NO:2) was subcloned into the pET14b expression vector (Novagen) and expressed in *Escherichia coli* BL21(DE3) cells. Uniformly  $^{15}\text{N}$ - and  $^{15}\text{N}/^{13}\text{C}$ -labelled proteins were prepared by growing bacteria in a minimal medium containing  $^{15}\text{NH}_4\text{Cl}$  with or without  $^{13}\text{C}_6$ -glucose. A uniformly  $^{15}\text{N}/^{13}\text{C}$ -labelled and fractionally deuterated protein sample was prepared by growing the cells in 75%  $^2\text{H}_2\text{O}$ . The bromodomain was purified by affinity chromatography on a nickel-IDA column (Invitrogen) followed by the removal of poly-His tag by thrombin cleavage. The final purification of the protein was achieved by size-exclusion chromatography. The acetyl-lysine-containing peptides were prepared on a MilliGen 9050 peptide synthesizer (Perkin Elmer) using Fmoc/HBTU chemistry. Acetyl-lysine was incorporated using the reagent Fmoc-Ac-Lys with HBTU/DIPEA activation. NMR samples contained approximately 1 mM protein in 100mM phosphate buffer of pH 6.5 and 5mM perdeuterated DTT and 0.5mM EDTA in  $\text{H}_2\text{O}/^2\text{H}_2\text{O}$  (9/1) or  $^2\text{H}_2\text{O}$ .

*NMR spectroscopy:* All NMR spectra were acquired at 30°C on a Bruker DRX600 or DRX500 spectrometer. The backbone assignments of the  $^1\text{H}$ ,  $^{13}\text{C}$ , and  $^{15}\text{N}$  resonances were achieved using deuterium-decoupled triple-resonance experiments of HNCACB and HN(CO)CACB [Yamazaki *et al.*, *J. Am. Chem. Soc.* **116**:11655-11666 (1994)] recorded using the uniformly  $^{15}\text{N}/^{13}\text{C}$ -labeled and fractionally deuterated protein. The



side-chain atoms were assigned from 3D HCCH-TOCSY [Clore and Gronenborn, *Meth. Enzymol.* **239**:249-363 (1994)] and (H)C(CO)NH-TOCSY [Logan *et al.*, *J. Biomol. NMR* **3**:225-231 (1993)] data collected on the uniformly  $^{15}\text{N}/^{13}\text{C}$ -labeled protein. Stereospecific assignments of methyl groups of the Val and Leu residues were  
 5 obtained using a fractionally  $^{13}\text{C}$ -labeled sample [Neri *et al.*, *Biochemistry* **28**:7510-7516 (1989)]. The NOE-derived distance restraints were obtained from  $^{15}\text{N}$ - or  $^{13}\text{C}$ -edited 3D NOESY spectra.  $\phi$ -angle restraints were determined based on the  $^3J_{\text{HN,H}\alpha}$  coupling constants measured in a 3D HNHA spectrum [Clore and Gronenborn, *Meth. Enzymol.* **239**:249-363 (1994)]. Slowly exchanging amide protons were  
 10 identified from a series of 2D  $^{15}\text{N}$ -HSQC spectra recorded after the  $\text{H}_2\text{O}$  buffer was changed to a  $^2\text{H}_2\text{O}$  buffer. The intermolecular NOEs used in defining the structure of the bromodomain/Ac-histamine complex were detected in  $^{13}\text{C}$ -edited ( $F_1$ ),  $^{13}\text{C}/^{15}\text{N}$ -filtered ( $F_3$ ) 3D NOESY spectrum [Clore and Gronenborn, *Meth. Enzymol.* **239**:249-363 (1994)]. All NMR spectra were processed with the NMRPipe/NMRDraw  
 15 programs and analyzed using NMRView [Johnson and Blevins, *J. Biomol., NMR* **4**:603-614 (1994)].

*Structure calculations:* Structures of the bromodomain were calculated with a distance geometry/simulated annealing protocol using the X-PLOR program [Brunger, A. *X-PLOR Version 3.1: A system for X-Ray crystallography and NMR*, Yale University Press, New Haven, CT, (1993)]. A total of 1324 manually assigned NOE-derived  
 20 distance restraints were obtained from the  $^{15}\text{N}$ - and  $^{13}\text{C}$ -edited NOE spectra. Further analysis of the NOE spectra was carried out by the iterative automated assignment procedure using ARIA [Nilges and O'Donoghue, *Prog. NMR Spectroscopy* **32**:107-139  
 25 (1998)], which integrates with X-PLOR for structure calculations. A total of 1519 unambiguous and 590 ambiguous distance restraints were identified from the NOE data by ARIA, many of which were checked and confirmed manually. The ARIA-assigned distance restraints were in agreement with the structures calculated using only the manually assigned NOE distance restraints, 28 hydrogen-bond distance  
 30 restraints for 14 hydrogen bonds, and 54  $\phi$ -angle restraints. The final structure calculations employed a total of 3515 NMR experimental restraints obtained from the manual and the ARIA-assisted assignments, 2843 of which were unambiguously assigned NOE-derived distance restraints that comprise of 1077 intra-residue, 621

sequential, 550 medium-range, and 595 long-range NOEs. For the ensemble of the final 30 structures, no distance and torsional angle restraints were violated by more than 0.3 Å and 5°, respectively. The total, distance violation, and dihedral violation energies were  $178.7 \pm 2.4$  kcal mol<sup>-1</sup>,  $41.6 \pm 0.9$  kcal mol<sup>-1</sup>, and  $0.50 \pm 0.06$  kcal mol<sup>-1</sup>, respectively. The Lennard-Jones potential which was not used during any refinement stage, was  $-526.2 \pm 16.8$  kcal mol<sup>-1</sup> for the final structures. Ramachandran plot analysis of the final structures (residues 727-828) with Procheck-NMR [Laskowski *et al.*, *J. Biomol. NMR* 8:477-486 (1996)] showed that  $71.0 \pm 0.6\%$ ,  $23.8 \pm 0.6\%$ ,  $3.5 \pm 0.2\%$ , and  $1.7 \pm 0.2\%$  of the non-Gly and non-Pro residues were in the most favorable, additionally allowed, generously allowed, and disallowed regions, respectively. The corresponding values for the residues in the four  $\alpha$ -helices (residues 727-743, 770-776, 785-802, and 807-827) were  $88.9 \pm 0.4\%$ ,  $11.0 \pm 0.4\%$ ,  $0.1 \pm 0.1\%$ , and  $0.0 \pm 0.0\%$ , respectively. The structure of the bromodomain/acetyl-histamine complex was determined using the free form structure and additional 25 intermolecular and 5 intra-ligand NOE-derived distance restraints.

*Site-directed mutagenesis:* Mutant proteins were prepared using the QuickChange site-directed mutagenesis kit (Stratagene). The presence of appropriate mutations was confirmed by DNA sequencing.

*Ligand titration:* Ligand titration experiments were performed by recording a series of 2D <sup>15</sup>N- and <sup>13</sup>C-HSQC spectra on the uniformly <sup>15</sup>N-, and <sup>15</sup>N/<sup>13</sup>C-labelled bromodomain (~0.3mM), respectively, in the presence of different amounts of ligand concentration ranging from 0 to approximately 2.0 mM. The protein sample and the stock solutions of the ligands were all prepared in the same aqueous buffer containing 100mM phosphate and 5mM perdeuterated DTT at pH 6.5.

The full length nucleic acid sequence of the human p300/CBP-associated factor (P/CAF) was obtained from GenBank. Accession No: U57317.2 (SEQ ID NO:1) :

```

1  gggggcgcgt  cgacgcggaa  aagaggccgt  ggggggcctc  ccagcgctgg  cagacaccgt
61  gaggctggca  gccgccggca  cgcacacctt  gtccgcagtc  ccgaggaaca  tgtccgcagc
121  cagggcgcg  agcagagtcc  cgggcaggag  aaccaaggga  gggcgtgtgc  tgtggcgggc
181  gcggcagcgg  cagcggagcc  gctagtcccc  tccctcctgg  gggagcagct  gccgccgctg
241  ccgccgccgc  caccaccatc  agcgcgcggg  gcccggccag  agcgagccgg  gcgagcggcg

```



2941 gattaattga caagtgattt tttttccccc tctgcttctt agaaactcac caagcagtgt  
3001 gcctaaagca aggt

The full length protein sequence of the human p300/CBP-associated factor (P/CAF)

5 was obtained from GenBank. Accession No: U57317.2, (SEQ ID NO:2):

1 MSEAGGAGPG GCGAGAGAGA GPGALPPQPA ALPPAPPQGS PCAAAGGSG ACGPATAVAA  
61 AGTAEGPGGG GSARIAVKA QLRAPRAKK LEKLGVSAC KAEESCKNG WKNPNPSPTP  
121 PRADLQQIIV SLTESCRSCS HALAAHVSHL ENVSEEMNR LLGIVLDVEY LFTCVHKEED  
181 ADTKQVYFYL FKLLRKSILQ RGKPVVEGSL EKKPPFEKPS IEQGVNNFVQ YKFSHLPAGE  
10 241 RQTIVELAKM FLNRINYWHL EAPSQRRLRS PNDDISGYKE NYTRWLCYCN VPQFCDLPR  
301 YETTQVFGRT LLRSVFTVMR RQLLEQARQE KDKLPLEKRT LILTHFPKFL SMLEEEVYSQ  
361 NSPIWDQDFL SASSRTSQLG IQTVINPPPV AGTISYNSTS SSLEQPNAGS SSPACKASSG  
421 LEANPGEKRR MTDSHVLEEA KKPRVMGDIP MELINEVMST ITDPAAMLGP ETNFLSAHSA  
481 RDEAARLEER RGVIEFHVVG NSLNQKPNKK ILMWLVLQON VFSHQLPRMP KEYITRLVFD  
15 541 PKHKTALIK DGRVIGGICF RMFPSQGFTE IVFCAVTSNE QVKGYGTHLM NHLKEYHIKH  
601 DILNFLTLYAD EYAIGYFKKQ GFSKEIKIPK TKYVGVIKDY EGATLMGCEL NPRIPTYEFS  
661 VIIKKQKEII KKLIERKQAG IRKVYPGLSC FKDGVRIPI ESIPGIRETG WKPSGKEKSK  
721 EPRDPDQLYS TLKSILQQVK SHQSAWPFME PVKRTEAPGY YEVIRFPMDL KTMSERLKNR  
781 YYVSKKLFMA DLQRVFTNCK EYNAAESEYY KANILEKFF FSKIKEAGLI DK  
20

## Results

The P/CAF bromodomain represents an extensive family of bromodomains (Figure 1). A large number of long-range nuclear Overhauser enhancement (NOE)-derived  
25 distance restraints were identified in the NMR data of the P/CAF bromodomain, yielding a well-defined three-dimensional structure (Figures 2A -2D). Table 1 shows the NMR chemical shift assignment of the P/CAF bromodomain. Table 2 shows the Unambiguous NOE-derived distance restraints. Table 3 shows the Ambiguous NOE-derived distance restraints. Table 4 shows the Hydrogen bond restraints. The NMR  
30 structure coordinates of the P/CAF bromodomain in the free and complexed to acetyl-histamine are shown in Tables 5 and 6, respectively.

The structure consists of a four-helix bundle (helices  $\alpha_Z$ ,  $\alpha_A$ ,  $\alpha_B$ , and  $\alpha_C$ ) with a left-handed twist, and a long intervening loop between helices  $\alpha_Z$  and  $\alpha_A$  (termed the  
35 ZA loop, Figure 2E). The four amphipathic  $\alpha$ -helices are packed tightly against one another in an antiparallel manner, with crossing angles for adjacent helices of  $\sim 16$ - $20^\circ$ . The up-and-down four-helix bundle can adapt two topological folds with opposite

handedness (Figures 2F-2G). The right-handed four-helix bundle fold occurs more commonly and is seen in proteins such as hemerythrin and cytochrome  $b_{562}$ . The left-handed fold of the bromodomain structure is less common, but also observed in proteins such as cytochrome  $b_5$  and T4 lysozyme [Richardson, J., *Adv. Protein Chem.*, **34**:167-339 (1989); Presnell and Cohen, *Proc. Natl. Acad. Sci. USA* **86**:6592-6596 (1989)]. This topological difference arises from the orientation of the loop between the first two helices (Fig. 2F-2G). The right-handed four-helix bundle proteins have a relatively short hairpin-like connection between the first two helices, which makes the “preferred” turn to the right at the top of the first helix [Richardson, J., *Adv. Protein Chem.*, **34**:167-339 (1989); Presnell and Cohen, *Proc. Natl. Acad. Sci. USA* **86**:6592-6596 (1989); Weber and Salemme, *Nature* **287**:82-84 (1980)]. In contrast, proteins with the left-handed fold usually have a long loop after the first helix and often contain additional secondary structural elements at the base of the helix bundle [Richardson, J., *Adv. Protein Chem.*, **34**:167-339 (1989); Presnell and Cohen, *Proc. Natl. Acad. Sci. USA* **86**:6592-6596 (1989)]. In the bromodomain structure, this long ZA loop has a defined conformation and is packed against the loop between helices  $\alpha_b$  and  $\alpha_c$  (termed the BC loop) to form a hydrophobic pocket. These tertiary interactions between the two loops appear to favor the left turn of the ZA loop, resulting in the left-handed four-helix bundle fold of the bromodomain. The hydrophobic pocket formed by loops ZA and BC is lined by residues Val752, Ala757, Tyr760, Val763, Tyr802 and Tyr809 (Fig. 2H), and appears to be a site for protein-protein interactions (see below). The pocket is located at one end of the four-helix bundle, opposite to the N- and C-termini of the protein. Interestingly, the ZA loop varies in length amongst different bromodomains, but almost always contains residues corresponding to Phe748, Pro751, Pro758, Tyr760, and Pro767 (Figure 1). The conservation of these residues within the ZA loop as well as residues within the  $\alpha$ -helical regions implies a similar left-handed four-helix bundle structure for the large family of bromodomains (Fig. 1).

The modular bromodomain structure supports the idea that bromodomain can act as a functional unit for protein-protein interactions. The observation that bromodomains are found in nearly all known nuclear HATs (A-type) that are known to promote transcription-related acetylation of histones on specific lysine residues, but not present in cytoplasmic HATs (B-type), prompted the determination of whether bromodomains

can interact with acetyl-lysine (AcK). The NMR titration of the P/CAF bromodomain were performed with a peptide (SGRGKGG-AcK-GLGK) derived from histone H4, in which Lys8 is acetylated (Lys8 is the major acetylation site in H4 for GCN5, a yeast homologue of P/CAF). Remarkably, the bromodomain could indeed bind the AcK peptide. Moreover, this interaction appeared to be specific, based on the  $^{15}\text{N}$ -HSQC spectra which showed that only a limited number of residues underwent chemical shift changes as a function of peptide concentration (Figure 3A). Conversely, the NMR titration of the bromodomain with a non-acetylated, but otherwise identical H4 peptide, showed no noticeable chemical shift changes, demonstrating that the interaction between the bromodomain and the lysine-acetylated H4 peptide was dependent upon acetylation of lysine. The dissociation constant ( $K_D$ ) for the AcK peptide was estimated to be  $346 \pm 54 \mu\text{M}$ . This binding is likely reinforced through additional interactions between bromodomain-containing proteins and target proteins. Notably, many chromatin-associated proteins contain two or multiple bromodomains (Figure 1). Indeed, binding with another lysine-acetylated peptide (RKSTGG-AcK-APRKQ) derived from the major acetylation site on histone H3 (residues 9-20) was also observed. Together, these data demonstrate that the P/CAF bromodomain has the ability to bind AcK peptides in an acetylation dependent manner.

Intriguingly, the bromodomain residues that exhibited the most significant  $^1\text{H}$  and  $^{15}\text{N}$  chemical shift changes on peptide binding are located near the hydrophobic pocket between the ZA and BC loops (Figure 3B). Because a similar pattern of amide chemical shift changes was observed with the two different AcK-containing peptides, it was surmised that the hydrophobic cavity is the primary binding site for AcK. This hypothesis was further supported by titration with acetyl-histamine, which mimics the chemical structure of the AcK side-chain (Figure 3C). Both  $^{15}\text{N}$ - and  $^{13}\text{C}$ -HSQC spectra showed that interaction with acetyl-histamine was also acetylation-dependent, involving the same set of residues that showed chemical shift perturbations with similar concentration dependence. It should be noted that the bromodomain did not bind to the amino acids acetyl-lysine or acetyl-histidine alone, possibly due to the presence of the charged amino, carboxyl, or carboxylate group adjacent to the acetyl moiety (Figure 3C). Taken together, these results strongly suggest that the P/CAF

bromodomain can interact with acetyl-lysine-containing proteins in a specific manner, and that this interaction is localized to the bromodomain hydrophobic cavity.

- To identify the key residues involved in bromodomain-AcK recognition, the NMR structure of the P/CAF bromodomain in complex with acetyl-histamine was elucidated. As anticipated, the acetylated moiety binds in the bromodomain hydrophobic pocket (Figure 4). The intermolecular interactions are largely hydrophobic in nature, with the methyl group of acetyl-histamine making extensive contacts with the side-chains of Val752, Ala757, and Tyr760, and the methylene groups of acetyl-histamine displaying specific NOEs to Val752, Ala757, Tyr760, Tyr802, and Tyr809. No intermolecular NOEs were observed for the imidazole ring of acetyl-histamine. From the spectral analysis it is clear that the structure of the bromodomain is very similar in both the free and complex forms.
- It is worth noting that the bromodomain-AcK recognition is reminiscent of the interactions between the histone acetyltransferase Hat1 and acetyl-CoA. Although the binding pockets of these two otherwise structurally unrelated proteins are composed of different secondary structural elements, the nature of acetyl-lysine recognition has striking similarities. In particular, Tyr809, Tyr802, Tyr760, and Val752 in the bromodomain appear to be related to Phe220, Phe261, Val254, and Ile217 of Hat1, respectively, in their interactions with the acetyl moiety. This observation may suggest an evolutionary convergent mechanism of acetyl-lysine recognition between bromodomains and histone acetyltransferases.
- To determine the relative contributions of residues within the hydrophobic cavity in bromodomain-AcK binding, site-directed mutagenesis was used to alter residues Tyr809, Tyr802, Tyr760, and Val752 (Table 7).

**Table 7. Structural and Functional Analysis of the P/CAF Bromodomain Mutants**

5	Bromodomain Proteins	Structural Integrity <sup>a</sup>	H4 AcK-Peptide Binding $K_D$ ( $\mu$ M) <sup>b</sup>
	Wild-Type	++++	346 $\pm$ 54
	Tyr809Ala	++++	No Binding <sup>c</sup>
10	Tyr802Ala	+++	> 10,000 <sup>d</sup>
	Tyr760Ala	+++	> 10,000
15	Val752Ala	++	> 10,000

a. The effects of mutations on the structural integrity of the bromodomain were assessed by using the <sup>15</sup>N-HSQC spectra. The amide <sup>1</sup>H/<sup>15</sup>N resonances of the mutant proteins were compared to those of the wild-type bromodomain to determine if the particular mutations lead to global or local structure disruption. Severe line-broadening of the amide resonances would indicate protein conformational exchange due to a decrease of structure stability resulting from point mutations. Structural integrity of the mutant proteins is expressed here relative to that of the wild-type, using the signs of “++++” for as stable as the wild-type, “+++” for mildly destabilized, “++” for moderately destabilized, and “-” for completely unfolded.

b. The ligand binding affinity ( $K_D$ ) of the bromodomain proteins was estimated by following chemical shift changes of amide peaks in the <sup>15</sup>N-HSQC spectra as a function of the ligand concentration.

c. No detectable ligand binding observed in the NMR titration.

d. Ligand binding affinity was significantly reduced and beyond the limit for reliable measurements by NMR titration.



Substitution of Ala for Tyr809 completely abrogated the bromodomain binding to the lysine-acetylated H4 peptide, while the Tyr802Ala, Tyr760Ala, and Val752Ala mutants had significantly reduced ligand binding affinity. To assess whether these mutations disrupted the overall bromodomain fold, the  $^{15}\text{N}$ -HSQC spectra of the mutants was compared to that of the wild-type protein. For the Tyr809Ala mutant, the amide chemical shifts were only affected for a few residues near the mutation site. However, mutations of the other residues in the hydrophobic binding pocket perturbed the local protein conformation to greater extents, particularly the ZA loop (Table 7). Thus, the NMR structural analysis and the mutagenesis studies show that Tyr809, which is structurally supported by Trp746 and Asn803 (Figure 4), is essential for the bromodomain interaction with the acetyl group of acetyl-lysine, while residues of Tyr802, Tyr760, and Val752 likely play both structural and functional roles in the recognition. These residues are highly conserved throughout the bromodomain family (Figure 1), suggesting that recognition of acetyl-lysine may be a feature of bromodomains, in general. Therefore, Val752, Ala757, Tyr760, Tyr802, Asn803, and Tyr809 are key amino acid residues for the P/CAF bromodomain binding to acetyl-lysine.

**Table 8: Amino Acid Sequences of Bromodomains Identified in Figure 1**

PROTEIN BD	SEQ ID NO:	GenBank Acc. No.	PROTEIN BD	SEQ ID NO:	GenBank Acc. No.
hsp/CAF	7	U57317	dmFSH-2	25	
5 hsGCN5	8	U57136	scBDF1-2	26	
ttP55	9	U47321	hsBR140	27	JC2069
scGCN5	10	Q03330	hsSMAP	28	X87613
hsP300	11	A54277	ggPB1-1	29	X90849
hsCBP	12	S39162	ggPB1-2	30	
10 mmCBP	13	S39161	ggPB1-3	31	
ceYNJ1	14	P34545	ggPB1-4	32	
hsCCG1-1	15	P21675	ggPB1-5	33	
msCCG1-1	16	D26114	spBRO-1	34	S54260
hsCCG1-2	17		spBRO-2	35	
15 msCCG1-2	18		hsSNF2a	36	S45251
hsRing3-1	19	P25440	hsBRG1	37	S39039
hsORFX-1	20	D26362	ggBRM	38	X91638
dmFSH-1	21	P13709	ggBRG1	39	X91637
scBDF1-1	22	P35817	hsTIF1b	40	X97548
20 hsRing3-2	23		mmTIF1b	41	X99644
hsORFX-2	24		mmTIF1a	42	S78219

The present invention is not to be limited in scope by the specific embodiments described herein. Indeed, various modifications of the invention in addition to those described herein will become apparent to those skilled in the art from the foregoing description and the accompanying figures. Such modifications are intended to fall within the scope of the appended claims.

It is further to be understood that all base sizes or amino acid sizes, and all molecular weight or molecular mass values, given for nucleic acids or polypeptides are approximate, and are provided for description.

- 5 Various publications are cited herein, the disclosures of which are hereby incorporated by reference herein in their entireties.

WHAT IS CLAIMED IS:

- 1 1. An isolated nucleic acid encoding a peptide consisting of about 21 to 40  
2 amino acids comprising a ZA loop of a bromodomain comprising the amino acid  
3 sequence of SEQ ID NO:3.
- 1 2. The isolated nucleic acid of Claim 1 further comprising a heterologous  
2 nucleotide sequence.
- 1 3. An isolated nucleic acid encoding a peptide consisting of about 21 to 40  
2 amino acids comprising a ZA loop of a bromodomain, wherein the bromodomain has  
3 an amino acid sequence selected from the group consisting of SEQ ID NOs. 7, 8, 9,  
4 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32,  
5 33, 34, 35, 36, 37, 38, 39, 40, 41, and 42.
- 1 4. The isolated nucleic acid of Claim 3 further comprising a heterologous  
2 nucleotide sequence.
- 1 5. A peptide consisting of about 21 to 40 amino acids comprising a ZA loop of  
2 a bromodomain comprising the amino acid sequence of SEQ ID NO:3.
- 1 6. A fusion protein or peptide comprising the peptide of Claim 5.
- 1 7. A peptide consisting of about 21 to 40 amino acids comprising a ZA loop of  
2 a bromodomain, wherein the bromodomain has an amino acid sequence selected from  
3 the group consisting of SEQ ID NOs. 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20,  
4 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, and  
5 42.

1 8. A fusion protein or peptide comprising the peptide of Claim 7.

1 9. An antibody raised against the peptide of Claim 7 or raised against an  
2 antigenic fragment thereof.

1 10. An antibody raised against the peptide of Claim 5.

1 11. A method of identifying a compound that modulates the affinity of a  
2 bromodomain for a ligand that comprises an acetyl-lysine,  
3 said method comprising:

4 (a) contacting the bromodomain and the ligand in the presence of the  
5 compound, wherein the bromodomain and the ligand bind in the absence of the  
6 compound; and

7 (b) measuring the affinity of the bromodomain for the ligand; wherein  
8 a compound is identified as a compound that modulates the affinity of the  
9 bromodomain for the ligand when there is a change in the affinity of the  
10 bromodomain for the ligand in the presence of the compound.

1 12. The method of Claim 11, wherein the affinity of the bromodomain for the  
2 ligand increases in the presence of the compound and wherein the compound is  
3 identified as a bromodomain-ligand complex promoting agent.

1 13. The method of Claim 11, wherein the affinity of the bromodomain for the  
2 ligand decreases in the presence of the compound and the compound is identified as an  
3 inhibitor.

1 14. The method of Claim 11, wherein the compound is selected by performing  
2 rational drug design with the set of atomic coordinates obtained from one or more of

3 Tables 1-6, wherein said selecting is performed in conjunction with computer  
4 modeling.

1 15. The method of Claim 11, wherein the compound is selected by performing  
2 rational drug design with the set of atomic coordinates obtained from a set of atomic  
3 coordinates defining the three-dimensional structure of a bromodomain consisting of  
4 the amino acid sequence of SEQ ID NO:7, wherein said selecting is performed in  
5 conjunction with computer modeling.

1 16. A method of identifying a compound that modulates the stability of a  
2 bromodomain-acetyl-lysine binding complex comprising:  
3 (a) contacting the bromodomain-acetyl-lysine binding complex in the  
4 presence of the compound wherein the bromodomain-acetyl-lysine binding complex  
5 forms in the absence of the compound; and  
6 (c) measuring the stability of the bromodomain-acetyl-lysine binding  
7 complex; wherein a compound is identified as a compound that modulates the stability  
8 of the bromodomain-acetyl-lysine binding complex, when there is a change in the  
9 stability of the bromodomain-acetyl-lysine binding complex in the presence of the  
10 compound.

1 17. The method of Claim 16, wherein the stability of the bromodomain-acetyl-  
2 lysine binding complex increases in the presence of the compound and wherein the  
3 compound is identified as a stabilizing agent.

1 18. The method of Claim 16, wherein the stability of the bromodomain-acetyl-  
2 lysine binding complex decreases in the presence of the compound and the compound  
3 is identified as an inhibitor.

1 19. The method of Claim 16, wherein the compound is selected by performing  
2 rational drug design with the set of atomic coordinates obtained from one or more of  
3 Tables 1-6, wherein said selecting is performed in conjunction with computer  
4 modeling.

1 20. The method of Claim 16, wherein the compound is selected by performing  
2 rational drug design with the set of atomic coordinates obtained from a set of atomic  
3 coordinates defining the three-dimensional structure of a bromodomain consisting of  
4 the amino acid sequence of SEQ ID NO:7, wherein said selecting is performed in  
5 conjunction with computer modeling.

1 21. A method of identifying a binding partner for a protein that comprises an  
2 acetyl-lysine said method comprising:

3 (a) contacting the protein with a polypeptide comprising a  
4 bromodomain; and

5 (b) determining whether the polypeptide binds to the protein; wherein  
6 a binding partner for a protein is identified when polypeptide binds to the protein.

1 22. The method of Claim 21 wherein the bromodomain has an amino acid  
2 sequence from selected from the group consisting of SEQ ID NOs. 7, 8, 9, 10, 11, 12,  
3 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35,  
4 36, 37, 38, 39, 40, 41 and 42.

1 23. An agent that can inhibit the binding of a bromodomain with a protein  
2 comprising an acetyl-lysine selected from the group consisting of : ISYGR-AcK-  
3 KRRQRR (SEQ ID NO:4), ARKSTGG-AcK-APRKQL (SEQ ID NO:5) and  
4 QSTSRHK-AcK-LMFKTE (SEQ ID NO:6).

002220-4-020350

## ABSTRACT OF THE INVENTION

5



Table 1

**NMR Chemical  
Shift Assignment  
of the P/CAF  
Bromodomain**

RES_ID	715	HETEROGENEITY	100	CA	62.320000	CG1	28.733000
RES_TYPE	GLY	N	121.192000	HA	4.038000	HG11	1.748000
SPIN_SYSTEM_ID	1	HN	8.416000	CB	38.640000	HG12	1.052000
HETEROGENEITY	100	CA	63.430000	HB1	3.211000	CG2	17.168000
END_RES_DEF		HA	4.331000	HB2	3.024000	HG2#	1.003000
		CB	30.930000	CD1	134.350000	CD1	13.863000
		HB1	1.815000	HD1	7.053000	HD1#	0.619000
		HB2	1.762000	CE1	119.481000	END_RES_DEF	
		CG	27.630000	HE1	6.882000		
		HG1	1.681000	END_RES_DEF		RES_ID	736
		CD	43.603000			RES_TYPE	LEU
		HD1	3.161000	RES_ID	730	SPIN_SYSTEM_ID	22
		END_RES_DEF		RES_TYPE	SER	HETEROGENEITY	100
RES_ID	716	RES_ID	724	SPIN_SYSTEM_ID	16	N	119.880000
RES_TYPE	SER	RES_TYPE	ASP	HETEROGENEITY	100	HN	8.841000
SPIN_SYSTEM_ID	2	SPIN_SYSTEM_ID	10	N	112.173000	CA	58.473000
HETEROGENEITY	100	HETEROGENEITY	100	HN	8.167000	HA	4.090000
END_RES_DEF		N	122.012000	HA	3.920000	CB	41.950000
		HN	8.273000	HB1	3.995000	HB1	2.090000
		CA	52.415000	END_RES_DEF		HB2	1.703000
		HA	4.874000	RES_ID	731	CG	27.330000
		CB	41.400000	RES_TYPE	THR	HG	1.759000
		HB1	2.754000	SPIN_SYSTEM_ID	17	CD1	26.530000
		HB2	2.692000	HETEROGENEITY	100	HD1#	1.061000
		END_RES_DEF		N	120.372000	CD2	23.776000
RES_ID	717	RES_ID	725	HN	8.059000	HD2#	0.977000
RES_TYPE	HIS	RES_TYPE	PRO	CA	66.730000	END_RES_DEF	
SPIN_SYSTEM_ID	3	SPIN_SYSTEM_ID	11	HA	3.924000	RES_ID	737
HETEROGENEITY	100	HETEROGENEITY	100	CB	68.930000	RES_TYPE	GLN
END_RES_DEF		CA	65.080000	HB	4.247000	SPIN_SYSTEM_ID	23
		HA	4.329000	CG2	21.570000	HETEROGENEITY	100
		CB	32.590000	HG2#	1.142000	N	117.256000
		HB1	2.326000	END_RES_DEF		HN	8.505000
		HB2	1.973000	RES_ID	732	CA	59.020000
		CG	27.632000	RES_TYPE	LEU	HA	4.032000
		HG1	2.028000	SPIN_SYSTEM_ID	18	CB	28.182000
		CD	51.310000	HETEROGENEITY	100	HB1	2.327000
		HD1	3.866000	N	120.536000	HB2	2.263000
		END_RES_DEF		HN	8.460000	CG	34.240000
RES_ID	719	RES_ID	726	CA	57.920000	HG1	2.536000
RES_TYPE	SER	RES_TYPE	ASP	HA	3.289000	HG2	2.461000
SPIN_SYSTEM_ID	5	SPIN_SYSTEM_ID	12	CB	39.750000	END_RES_DEF	
HETEROGENEITY	100	HETEROGENEITY	100	HB1	1.532000	RES_ID	738
END_RES_DEF		N	119.716000	HB2	0.294000	RES_TYPE	GLN
		HN	8.397000	CG	24.880000	SPIN_SYSTEM_ID	24
		CA	55.720000	HG	1.683000	HETEROGENEITY	100
		HA	4.692000	CD1	25.429000	N	118.896000
		CB	40.550000	HD1#	0.469000	HN	8.033000
		HB1	2.792000	CD2	19.921000	CA	59.574000
		HB2	2.730000	HD2#	-0.193000	HA	4.196000
		END_RES_DEF		END_RES_DEF		CB	29.835000
RES_ID	720	RES_ID	727	RES_ID	733	HB1	2.482000
RES_TYPE	LYS	RES_TYPE	GLN	RES_TYPE	LYS	HB2	2.469000
SPIN_SYSTEM_ID	6	SPIN_SYSTEM_ID	13	SPIN_SYSTEM_ID	19	CG	35.342000
HETEROGENEITY	100	HETEROGENEITY	100	HETEROGENEITY	100	HG1	2.840000
CA	56.296000	N	121.356000	N	118.568000	HG2	2.467000
HA	4.361000	HN	8.196000	HN	8.563000	NE2	110.369000
CB	33.140000	CA	55.920000	CA	60.125000	HE21	7.022000
HB1	1.882000	HA	4.163000	HA	3.679000	HE22	6.916000
HB2	1.684000	CB	28.730000	CB	32.588000	END_RES_DEF	
CG	25.430000	HB1	2.148000	HB1	1.729000	RES_ID	739
HG1	1.585000	CG	34.240000	HB2	1.360000	RES_TYPE	VAL
HG2	1.433000	HG1	2.524000	CG	24.880000	SPIN_SYSTEM_ID	25
CD	29.834000	HG2	2.371000	HG1	1.280000	HETEROGENEITY	100
HD1	1.703000	END_RES_DEF		CD	29.835000	N	119.716000
CE	41.960000	RES_ID	728	HD1	1.585000	HN	8.526000
HE1	3.003000	RES_TYPE	LEU	CE	41.960000	CA	67.830000
END_RES_DEF		SPIN_SYSTEM_ID	14	HE1	2.918000	HA	3.844000
RES_ID	721	HETEROGENEITY	100	END_RES_DEF		CB	32.030000
RES_TYPE	GLU	N	121.356000	RES_ID	734	HB	2.384000
SPIN_SYSTEM_ID	7	HN	8.196000	RES_TYPE	SER	CG1	23.330000
HETEROGENEITY	100	CA	55.920000	SPIN_SYSTEM_ID	20	HG1#	1.183000
N	122.990000	HA	4.163000	HETEROGENEITY	100	CG2	22.120000
HN	8.317000	CB	28.730000	N	113.157000	HG2#	1.033000
CA	54.620000	HB1	1.847000	HN	7.540000	END_RES_DEF	
HA	4.540000	HB2	1.555000	CA	61.227000	RES_ID	740
CB	29.830000	CG	27.080000	HA	4.281000	RES_TYPE	LYS
HB1	2.024000	HG	1.480000	CB	63.879000	SPIN_SYSTEM_ID	26
HB2	1.893000	CD1	25.970000	HB1	4.060000	HETEROGENEITY	100
CG	35.893000	HD1#	0.794000	END_RES_DEF		N	114.633000
HG1	2.271000	CD2	23.226000	RES_ID	735	HN	8.572000
END_RES_DEF		HD2#	0.786000	RES_TYPE	ILE	CA	59.574000
RES_ID	722	END_RES_DEF		SPIN_SYSTEM_ID	21	HA	3.886000
RES_TYPE	PRO	RES_ID	729	HETEROGENEITY	100	CB	32.380000
SPIN_SYSTEM_ID	8	RES_TYPE	TYR	N	120.700000	HB1	1.873000
HETEROGENEITY	100	SPIN_SYSTEM_ID	15	HN	7.951000	HG1	1.022000
CA	63.430000	HETEROGENEITY	100	CA	65.080000	HD1	1.520000
HA	4.393000	N	119.060000	HA	3.786000	END_RES_DEF	
CB	32.030000	HN	8.021000	CB	38.095000	RES_ID	741
HB1	2.224000			HB	1.879000	RES_TYPE	SER
HB2	1.880000						
CG	27.630000						
HG1	2.028000						
CD	50.760000						
HD2	3.656000						
HD1	3.800000						
END_RES_DEF							
RES_ID	723						
RES_TYPE	ARG						
SPIN_SYSTEM_ID	9						

SPIN\_SYSTEM\_ID 27  
 HETEROGENEITY 100  
 N 110.369000  
 HN 7.557000  
 CA 59.024000  
 HA 4.448000  
 CB 63.980000  
 HB1 4.004000  
 END\_RES\_DEF

RES\_ID 742  
 RES\_TYPE HIS  
 SPIN\_SYSTEM\_ID 28  
 HETEROGENEITY 100

N 125.619000  
 HN 7.536000  
 CA 58.473000  
 HA 3.967000  
 CB 32.588000  
 HB1 2.990000  
 HB2 2.799000  
 CD2 118.930000  
 HD2 4.978000  
 CE1 138.755000  
 HE1 7.522000  
 END\_RES\_DEF

RES\_ID 743  
 RES\_TYPE GLN  
 SPIN\_SYSTEM\_ID 29  
 HETEROGENEITY 100

N 128.571000  
 HN 8.543000  
 CA 59.125000  
 HA 4.209000  
 CB 29.834000  
 HB1 2.111000  
 CG 33.690000  
 HG1 2.390000  
 NE2 112.173000  
 HE21 7.581000  
 HE22 6.870000  
 END\_RES\_DEF

RES\_ID 744  
 RES\_TYPE SER  
 SPIN\_SYSTEM\_ID 30  
 HETEROGENEITY 100

N 119.060000  
 HN 11.668000  
 CA 60.125000  
 HA 4.838000  
 CB 63.980000  
 HB1 4.334000  
 HB2 3.926000  
 END\_RES\_DEF

RES\_ID 745  
 RES\_TYPE ALA  
 SPIN\_SYSTEM\_ID 31  
 HETEROGENEITY 100

N 117.584000  
 HN 7.868000  
 CA 53.510000  
 HA 4.396000  
 CB 20.470000  
 HB# 1.688000  
 END\_RES\_DEF

RES\_ID 746  
 RES\_TYPE TRP  
 SPIN\_SYSTEM\_ID 32  
 HETEROGENEITY 100

N 116.600000  
 HN 7.135000  
 CA 60.691000  
 HA 4.368000  
 CB 27.630000  
 HB1 3.594000  
 HB2 3.351000  
 CD1 128.843000  
 HD1 7.897000  
 NE1 110.861000  
 HE1 10.474000  
 CE3 122.234000  
 HE3 7.336000  
 CZ2 116.177000  
 HZ2 7.382000  
 CZ3 123.336000  
 HZ3 7.197000  
 CH2 126.089000  
 HH2 7.150000  
 END\_RES\_DEF

RES\_ID 747

RES\_TYPE PRO  
 SPIN\_SYSTEM\_ID 33  
 HETEROGENEITY 100

CA 64.531000  
 HA 3.756000  
 CB 29.835000  
 HB1 0.487000  
 HB2 -0.783000  
 CG 26.530000  
 HG1 0.233000  
 HG2 -0.931000  
 CD 50.212000  
 HD2 1.567000  
 HD1 2.177000  
 END\_RES\_DEF

RES\_ID 748  
 RES\_TYPE PHE  
 SPIN\_SYSTEM\_ID 34  
 HETEROGENEITY 100

N 113.321000  
 HN 7.585000  
 CA 55.719000  
 HA 4.930000  
 CB 39.202000  
 HB1 3.491000  
 HB2 2.532000  
 CD1 133.248000  
 HD1 7.099000  
 HE1 7.174000  
 HZ 7.296000  
 END\_RES\_DEF

RES\_ID 749  
 RES\_TYPE MET  
 SPIN\_SYSTEM\_ID 35  
 HETEROGENEITY 100

N 117.748000  
 HN 7.115000  
 CA 56.820000  
 HA 4.286000  
 CB 32.590000  
 HB1 2.233000  
 HB2 2.174000  
 CG 33.140000  
 HG1 2.851000  
 CE 17.168000  
 HE# 2.175000  
 END\_RES\_DEF

RES\_ID 750  
 RES\_TYPE GLU  
 SPIN\_SYSTEM\_ID 36  
 HETEROGENEITY 100

N 113.813000  
 HN 7.709000  
 CA 53.516000  
 HA 4.849000  
 CB 31.487000  
 HB1 2.091000  
 HB2 1.730000  
 CG 35.893000  
 HG1 2.164000  
 END\_RES\_DEF

RES\_ID 751  
 RES\_TYPE PRO  
 SPIN\_SYSTEM\_ID 37  
 HETEROGENEITY 100

CA 62.879000  
 HA 4.242000  
 CB 32.040000  
 HB1 2.328000  
 HB2 1.683000  
 CG 27.080000  
 HG1 2.126000  
 HG2 1.978000  
 CD 50.763000  
 HD1 3.670000  
 END\_RES\_DEF

RES\_ID 752  
 RES\_TYPE VAL  
 SPIN\_SYSTEM\_ID 38  
 HETEROGENEITY 100

N 124.450000  
 HN 8.124000  
 CA 63.430000  
 HA 3.553000  
 CB 32.580000  
 HB 1.145000  
 CG1 21.573000  
 HG1# 0.464000  
 CG2 21.573000  
 HG2# 0.169000

END\_RES\_DEF

RES\_ID 753  
 RES\_TYPE LYS  
 SPIN\_SYSTEM\_ID 39  
 HETEROGENEITY 100

N 129.883000  
 HN 9.045000  
 CA 56.310000  
 HA 4.370000  
 CB 32.880000  
 HB1 1.873000  
 HG1 1.435000  
 HD1 1.673000  
 HE1 2.985000  
 END\_RES\_DEF

RES\_ID 754  
 RES\_TYPE ARG  
 SPIN\_SYSTEM\_ID 40  
 HETEROGENEITY 100

N 120.208000  
 HN 8.054000  
 END\_RES\_DEF

RES\_ID 755  
 RES\_TYPE THR  
 SPIN\_SYSTEM\_ID 41  
 HETEROGENEITY 100

CA 63.430000  
 HA 4.038000  
 CB 68.380000  
 HB 4.293000  
 CG2 22.670000  
 HG2# 1.267000  
 END\_RES\_DEF

RES\_ID 756  
 RES\_TYPE GLU  
 SPIN\_SYSTEM\_ID 42  
 HETEROGENEITY 100

N 118.732000  
 HN 7.209000  
 CA 56.270000  
 HA 4.448000  
 CB 30.930000  
 HB1 2.174000  
 HB2 2.000000  
 CG 36.440000  
 HG1 2.292000  
 END\_RES\_DEF

RES\_ID 757  
 RES\_TYPE ALA  
 SPIN\_SYSTEM\_ID 43  
 HETEROGENEITY 100

N 122.504000  
 HN 7.379000  
 CA 50.220000  
 HA 4.937000  
 CB 19.370000  
 HB# 1.082000  
 END\_RES\_DEF

RES\_ID 758  
 RES\_TYPE PRO  
 SPIN\_SYSTEM\_ID 44  
 HETEROGENEITY 100

CA 65.080000  
 HA 4.496000  
 CB 31.487000  
 HB1 2.374000  
 HB2 2.027000  
 CG 27.632000  
 HG1 2.122000  
 HG2 2.038000  
 CD 50.212000  
 HD2 3.515000  
 HD1 3.717000  
 END\_RES\_DEF

RES\_ID 759  
 RES\_TYPE GLY  
 SPIN\_SYSTEM\_ID 45  
 HETEROGENEITY 100  
 END\_RES\_DEF

RES\_ID 760  
 RES\_TYPE TYR  
 SPIN\_SYSTEM\_ID 46  
 HETEROGENEITY 100  
 N 122.504000  
 HN 7.945000  
 CA 62.328000  
 HA 3.536000

CB 39.750000  
 HB1 2.689000  
 HB2 2.487000  
 CD1 133.799000  
 HD1 5.120000  
 CE1 118.379000  
 HE1 6.070000  
 END\_RES\_DEF

RES\_ID 761  
 RES\_TYPE TYR  
 SPIN\_SYSTEM\_ID 47  
 HETEROGENEITY 100

N 113.157000  
 HN 8.225000  
 CA 60.676000  
 HA 4.101000  
 CB 37.550000  
 HB1 3.189000  
 HB2 2.801000  
 CD1 134.901000  
 HD1 7.342000  
 CE1 118.930000  
 HE1 6.646000  
 END\_RES\_DEF

RES\_ID 762  
 RES\_TYPE GLU  
 SPIN\_SYSTEM\_ID 48  
 HETEROGENEITY 100

N 117.912000  
 HN 7.702000  
 CA 57.922000  
 HA 4.209000  
 CB 29.480000  
 HB1 2.086000  
 CG 37.545000  
 HG1 2.325000  
 HG2 2.265000  
 END\_RES\_DEF

RES\_ID 763  
 RES\_TYPE VAL  
 SPIN\_SYSTEM\_ID 49  
 HETEROGENEITY 100

N 115.453000  
 HN 7.135000  
 CA 63.430000  
 HA 4.077000  
 CB 33.690000  
 HB 2.015000  
 CG1 21.020000  
 HG1# 1.045000  
 CG2 21.574000  
 HG2# 0.991000  
 END\_RES\_DEF

RES\_ID 764  
 RES\_TYPE ILE  
 SPIN\_SYSTEM\_ID 50  
 HETEROGENEITY 100

N 122.832000  
 HN 7.947000  
 CA 57.920000  
 HA 3.916000  
 CB 34.240000  
 HB 1.205000  
 CG1 24.878000  
 HG11 0.798000  
 HG12 0.216000  
 CG2 16.617000  
 HG2# 0.380000  
 CD1 9.457000  
 HD1# 0.537000  
 END\_RES\_DEF

RES\_ID 765  
 RES\_TYPE ARG  
 SPIN\_SYSTEM\_ID 51  
 HETEROGENEITY 100

N 125.291000  
 HN 7.749000  
 CA 57.371000  
 HA 3.875000  
 CB 30.936000  
 HB1 1.388000  
 HB2 1.211000  
 CG 27.080000  
 HG1 1.319000  
 HG2 1.173000  
 CD 43.052000  
 HD1 2.971000  
 END\_RES\_DEF

RES\_ID 766

N	128.079000
HN	8.738000
CA	60.676000
HA	4.198000
CB	32.037000
HB1	2.330000
HB2	2.224000
CG	25.280000
HG1	1.483000
HG2	1.403000
CD	30.385000
HD1	1.793000
HD2	1.696000
CE	41.950000
HE1	2.965000

HETEROGENEITY		100
N	120.208000	
HN	8.856000	
CA	58.470000	
HA	4.691000	
CB	42.621000	
HB1	2.295000	
HB2	1.925000	
CG	27.080000	
HG	1.832000	

ND\_RES\_DEF

ES_ID	783
ES_TYPE	IND

RES_ID	789
RES_TYPE	MET
SPIN_SYSTEM_ID	75
HETEROGENEITY	100



HN 8.068000  
CA 56.270000  
HA 4.329000  
CB 38.646000  
HB1 2.877000  
HB2 2.834000  
END\_RES\_DEF

RES\_ID 815  
RES\_TYPE ILE  
SPIN\_SYSTEM\_ID 101  
HETEROGENEITY 100  
N 119.880000  
HN 7.912000  
CA 65.080000  
HA 3.646000  
CB 39.197000  
HB 1.924000  
CG1 29.284000  
HG11 1.882000  
HG12 1.201000  
CG2 17.718000  
HG2# 1.017000  
CD1 13.863000  
HD1# 0.940000  
END\_RES\_DEF

RES\_ID 816  
RES\_TYPE LEU  
SPIN\_SYSTEM\_ID 102  
HETEROGENEITY 100  
N 122.504000  
HN 8.556000  
CA 56.820000  
HA 3.670000  
CB 41.951000  
HB1 1.405000  
HB2 1.199000  
CG 26.530000  
HG 1.580000  
CD1 24.327000  
HD1# 0.701000  
CD2 25.429000  
HD2# 0.696000  
END\_RES\_DEF

RES\_ID 817  
RES\_TYPE GLU  
SPIN\_SYSTEM\_ID 103  
HETEROGENEITY 100  
N 120.700000  
HN 8.073000  
CA 60.125000  
HA 3.185000  
CB 29.835000  
HB1 1.720000  
HB2 1.310000  
CG 37.545000  
HG1 2.001000  
HG2 1.922000  
END\_RES\_DEF

RES\_ID 818  
RES\_TYPE LYS  
SPIN\_SYSTEM\_ID 104  
HETEROGENEITY 100  
N 117.584000  
HN 7.145000  
CA 59.688000  
HA 4.075000  
CB 32.588000  
HB1 1.929000  
CG 25.644000  
HG1 1.492000  
CD 29.284000  
HD1 1.681000  
CE 41.963000  
HE1 2.964000  
END\_RES\_DEF

RES\_ID 819  
RES\_TYPE PHE  
SPIN\_SYSTEM\_ID 105  
HETEROGENEITY 100  
N 121.028000  
HN 7.869000  
CA 61.230000  
HA 4.328000  
CB 39.200000  
HB1 3.133000  
HB2 3.047000  
CD1 133.800000  
HD1 7.180000  
END\_RES\_DEF

RES\_ID 820  
RES\_TYPE PHE  
SPIN\_SYSTEM\_ID 106  
HETEROGENEITY 100  
N 120.700000  
HN 9.126000  
CA 60.691000  
HA 3.961000  
CB 38.640000  
HB1 3.289000  
HB2 3.067000  
CD1 133.248000  
HD1 6.904000  
CE1 132.698000  
HE1 7.011000  
END\_RES\_DEF

RES\_ID 821  
RES\_TYPE PHE  
SPIN\_SYSTEM\_ID 107  
HETEROGENEITY 100  
N 118.076000  
HN 8.359000  
CA 61.770000  
HA 3.840000  
CB 38.090000  
HB1 3.064000  
CD1 133.248000  
HD1 7.175000  
CE1 132.698000  
HE1 7.294000  
CZ 131.596000  
HZ 7.430000  
END\_RES\_DEF

RES\_ID 822  
RES\_TYPE SER  
SPIN\_SYSTEM\_ID 108  
HETEROGENEITY 100  
N 114.961000  
HN 7.906000  
CA 61.773000  
HA 4.200000  
CB 62.879000  
HB1 4.007000  
END\_RES\_DEF

RES\_ID 823  
RES\_TYPE LYS  
SPIN\_SYSTEM\_ID 109  
HETEROGENEITY 100  
N 120.864000  
HN 7.938000  
CA 56.820000  
HA 4.008000  
CB 31.487000  
HB1 1.730000  
HB2 1.567000  
CG 23.226000  
HG1 0.833000  
CD 27.080000  
HD1 1.403000  
CE 42.501000  
HE1 2.569000  
HE2 2.422000  
END\_RES\_DEF

RES\_ID 824  
RES\_TYPE ILE  
SPIN\_SYSTEM\_ID 110  
HETEROGENEITY 100  
N 116.928000  
HN 8.101000  
CA 64.530000  
HA 3.818000  
CB 36.990000  
HB 1.746000  
CG1 26.530000  
HG11 1.140000  
HG12 1.073000  
CG2 18.820000  
HG2# 0.654000  
CD1 13.312000  
HD1# 0.541000  
END\_RES\_DEF

RES\_ID 825  
RES\_TYPE LYS  
SPIN\_SYSTEM\_ID 111  
HETEROGENEITY 100  
N 122.176000  
HN 7.546000  
CA 59.024000  
HA 4.043000  
CB 32.360000

HB1 1.879000  
HB2 1.757000  
CG 24.878000  
HG1 1.390000  
HG2 1.302000  
CD 29.284000  
HD1 1.633000  
CE 41.400000  
HE1 2.913000  
END\_RES\_DEF

RES\_ID 826  
RES\_TYPE GLU  
SPIN\_SYSTEM\_ID 112  
HETEROGENEITY 100  
N 121.192000  
HN 8.063000  
CA 59.024000  
HA 3.995000  
CB 29.834000  
HB1 2.058000  
CG 36.050000  
HG1 2.342000  
HG2 2.205000  
END\_RES\_DEF

RES\_ID 827  
RES\_TYPE ALA  
SPIN\_SYSTEM\_ID 113  
HETEROGENEITY 100  
N 117.748000  
HN 7.620000  
CA 52.410000  
HA 4.291000  
CB 19.920000  
HB# 1.358000  
END\_RES\_DEF

RES\_ID 828  
RES\_TYPE GLY  
SPIN\_SYSTEM\_ID 114  
HETEROGENEITY 100  
N 126.767000  
HN 7.744000  
CA 45.902000  
HA1 4.019000  
HA2 3.935000  
END\_RES\_DEF

RES\_ID 829  
RES\_TYPE LEU  
SPIN\_SYSTEM\_ID 115  
HETEROGENEITY 100  
N 117.912000  
HN 7.742000  
CA 55.719000  
HA 4.215000  
CB 43.052000  
HB1 1.562000  
CG 27.632000  
HG 1.536000  
CD1 23.776000  
HD1# 0.711000  
END\_RES\_DEF

RES\_ID 830  
RES\_TYPE ILE  
SPIN\_SYSTEM\_ID 116  
HETEROGENEITY 100  
N 115.453000  
HN 7.458000  
CA 60.676000  
HA 4.232000  
CB 39.748000  
HB 1.810000  
CG1 27.080000  
HG11 1.314000  
HG12 0.918000  
CG2 17.718000  
HG2# 0.815000  
CD1 13.312000  
HD1# 0.794000  
END\_RES\_DEF

RES\_ID 831  
RES\_TYPE ASP  
SPIN\_SYSTEM\_ID 117  
HETEROGENEITY 100  
N 123.488000  
HN 8.270000  
CA 54.620000  
HA 4.571000  
CB 41.400000  
HB1 2.693000  
HB2 2.540000

END\_RES\_DEF  
RES\_ID 832  
RES\_TYPE LYS  
SPIN\_SYSTEM\_ID 118  
HETEROGENEITY 100  
N 125.450000  
HN 7.774000  
CA 57.720000  
HA 4.082000  
CB 33.410000  
END\_RES\_DEF

Unambiguous NOE-derived Inter-proton Distance Restraints

[illegible]

(( segid "BPD" * and readid 106 and name HB2 ))	0.11000E+01 volume	0.11998E+04 ppm1	5.740 ppm2	3.674
I { 451 }	1.300 1.300 peak 431 weight			
(( segid "BPD" * and readid 107 and name HN ))				
I { 452 }	2.800 2.800 peak 431 weight			
(( segid "BPD" * and readid 106 and name HN ))				
I { 461 }	2.900 2.100 2.100 peak 451 weight	0.11000E+01 volume	0.34881E+03 ppm1	9.713
(( segid "BPD" * and readid 107 and name HN ))				
I { 462 }	3.100 3.100 2.000 peak 461 weight	0.11000E+01 volume	0.11255E+03 ppm1	7.783
(( segid "BPD" * and readid 107 and name HN ))				
I { 441 }	2.400 1.700 1.700 peak 471 weight	0.11000E+01 volume	0.57631E+03 ppm1	4.443
(( segid "BPD" * and readid 107 and name HN ))				
I { 501 }	2.600 2.600 2.300 peak 431 weight	0.11000E+01 volume	0.16715E+03 ppm1	8.960
(( segid "BPD" * and readid 106 and name HN ))				
I { 511 }	1.700 1.700 peak 501 weight	0.11000E+01 volume	0.57313E+03 ppm1	8.529
(( segid "BPD" * and readid 108 and name HN ))				
I { 700 }	1.800 1.800 peak 511 weight	0.11000E+01 volume	0.52916E+03 ppm1	8.526
(( segid "BPD" * and readid 108 and name HN ))				
I { 531 }	2.400 1.400 1.400 peak 521 weight	0.11000E+01 volume	0.97946E+03 ppm1	4.583
(( segid "BPD" * and readid 109 and name HN ))				
I { 541 }	1.300 1.300 peak 531 weight	0.11000E+01 volume	0.11922E+04 ppm1	4.615
(( segid "BPD" * and readid 110 and name HN ))				
I { 551 }	2.100 2.100 peak 541 weight	0.11000E+01 volume	0.32815E+03 ppm1	8.714
(( segid "BPD" * and readid 110 and name HN ))				
I { 561 }	2.800 2.000 2.000 peak 551 weight	0.11000E+01 volume	0.43619E+03 ppm1	4.451
(( segid "BPD" * and readid 110 and name HN ))				
I { 581 }	2.500 1.600 1.600 peak 561 weight	0.11000E+01 volume	0.77600E+03 ppm1	8.714
(( segid "BPD" * and readid 109 and name HN ))				
I { 591 }	2.800 2.000 2.000 peak 581 weight	0.11000E+01 volume	0.41842E+03 ppm1	8.572
(( segid "BPD" * and readid 111 and name HN ))				
I { 601 }	2.600 2.200 peak 591 weight	0.11000E+01 volume	0.17650E+03 ppm1	8.168
(( segid "BPD" * and readid 111 and name HN ))				
I { 621 }	2.500 1.600 1.600 peak 601 weight	0.11000E+01 volume	0.71549E+03 ppm1	4.650
(( segid "BPD" * and readid 110 and name HN ))				
I { 631 }	2.800 2.000 2.000 peak 621 weight	0.11000E+01 volume	0.38461E+03 ppm1	8.153
(( segid "BPD" * and readid 112 and name HN ))				
I { 641 }	1.300 1.300 peak 631 weight	0.11000E+01 volume	0.13944E+04 ppm1	4.607
(( segid "BPD" * and readid 112 and name HN ))				
I { 3400 }	2.900 2.100 peak 641 weight	0.11000E+01 volume	0.12113E+03 ppm1	8.658
(( segid "BPD" * and readid 112 and name HN ))				
I { 661 }	3.300 3.100 2.000 peak 651 weight	0.11000E+01 volume	0.96111E+02 ppm1	2.618
(( segid "BPD" * and readid 112 and name HN ))				
I { 671 }	1.200 1.200 peak 661 weight	0.11000E+01 volume	0.15417E+04 ppm1	8.668
(( segid "BPD" * and readid 113 and name HN ))				
I { 6800 }	2.000 2.000 2.000 peak 671 weight	0.11000E+01 volume	0.42978E+03 ppm1	8.217
(( segid "BPD" * and readid 113 and name HN ))				
I { 701 }	2.400 1.400 1.400 peak 681 weight	0.11000E+01 volume	0.99017E+03 ppm1	1.967
(( segid "BPD" * and readid 114 and name HN ))				
I { 721 }	2.800 2.000 2.000 peak 701 weight	0.11000E+01 volume	0.42792E+03 ppm1	8.376
(( segid "BPD" * and readid 114 and name HN ))				
I { 7300 }	1.600 1.600 peak 721 weight	0.11000E+01 volume	0.73856E+03 ppm1	6.219
				8.351

—















# D O C U M E N T





—

[illegible]

```

(( segid "BRD " and resid 52 and name HBI ))
ASSI { 7061 } 2 700 1 800 1 800 peak 7051 weight 0 11000E+01 volume 0 48728E+03 ppm1 9 004 ppm2 3 641
(( segid "BRD " and resid 52 and name HN ))
(( segid "BRD " and resid 52 and name H2 ))
ASSI { 7062 } 2 500 1 600 1 600 peak 7051 weight 0 11000E+01 volume 0 7250E+03 ppm1 9 003 ppm2 3 514
(( segid "BRD " and resid 52 and name HN ))
(( segid "BRD " and resid 52 and name H2 ))
ASSI { 7071 } 2 600 1 700 1 700 peak 7071 weight 0 11000E+01 volume 0 63507E+03 ppm1 9 004 ppm2 8 356
(( segid "BRD " and resid 52 and name HN ))
(( segid "BRD " and resid 52 and name H2 ))
ASSI { 7101 } 3 000 2 200 2 200 peak 7091 weight 0 11000E+01 volume 0 26640E+03 ppm1 9 004 ppm2 4 482
(( segid "BRD " and resid 52 and name HN ))
(( segid "BRD " and resid 52 and name H2 ))
ASSI { 7111 } 3 500 3 800 1 600 peak 7111 weight 0 11000E+01 volume 0 10363E+03 ppm1 9 004 ppm2 4 230
(( segid "BRD " and resid 52 and name HN ))
(( segid "BRD " and resid 52 and name H2 ))
ASSI { 7121 } 3 200 2 700 2 200 peak 7141 weight 0 11000E+01 volume 0 14988E+03 ppm1 8 496 ppm2 3 103
(( segid "BRD " and resid 52 and name HN ))
(( segid "BRD " and resid 52 and name H2 ))
ASSI { 7131 } 3 400 2 900 2 100 peak 7151 weight 0 11000E+01 volume 0 11956E+03 ppm1 8 498 ppm2 2 700
(( segid "BRD " and resid 52 and name HN ))
(( segid "BRD " and resid 52 and name H2 ))
ASSI { 7141 } 3 200 2 600 2 300 peak 7171 weight 0 11000E+01 volume 0 17218E+03 ppm1 8 499 ppm2 2 504
(( segid "BRD " and resid 52 and name HN ))
(( segid "BRD " and resid 52 and name H2 ))
ASSI { 7151 } 3 400 2 900 2 100 peak 7191 weight 0 11000E+01 volume 0 9473E+03 ppm1 9 359 ppm2 2 947
(( segid "BRD " and resid 52 and name HN ))
(( segid "BRD " and resid 52 and name H2 ))
ASSI { 7201 } 2 800 1 600 1 600 peak 7201 weight 0 11000E+01 volume 0 75718E+03 ppm1 9 358 ppm2 2 827
(( segid "BRD " and resid 52 and name HN ))
(( segid "BRD " and resid 52 and name H2 ))
ASSI { 7211 } 3 600 3 200 1 900 peak 7201 weight 0 11000E+01 volume 0 93147E+02 ppm1 9 359 ppm2 2 088
(( segid "BRD " and resid 52 and name HN ))
(( segid "BRD " and resid 52 and name H2 ))
ASSI { 7221 } 4 300 1 200 1 200 peak 7211 weight 0 11000E+01 volume 0 28921E+02 ppm1 9 358 ppm2 2 411
(( segid "BRD " and resid 52 and name HN ))
(( segid "BRD " and resid 52 and name H2 ))
ASSI { 7231 } 3 400 2 900 2 100 peak 7221 weight 0 11000E+01 volume 0 13021E+03 ppm1 9 359 ppm2 2 292
(( segid "BRD " and resid 52 and name HN ))
(( segid "BRD " and resid 52 and name H2 ))
ASSI { 7241 } 3 300 2 700 2 200 peak 7231 weight 0 11000E+01 volume 0 15703E+03 ppm1 9 195 ppm2 2 463
(( segid "BRD " and resid 52 and name HN ))
(( segid "BRD " and resid 52 and name H2 ))
ASSI { 7251 } 4 600 4 600 0 900 peak 7261 weight 0 11000E+01 volume 0 20621E+02 ppm1 8 885 ppm2 4 475
(( segid "BRD " and resid 52 and name HN ))
(( segid "BRD " and resid 52 and name H2 ))
ASSI { 7261 } 3 400 2 900 2 100 peak 7271 weight 0 11000E+01 volume 0 12458E+03 ppm1 8 883 ppm2 2 422
(( segid "BRD " and resid 52 and name HN ))
(( segid "BRD " and resid 52 and name H2 ))
ASSI { 7271 } 3 500 3 100 2 000 peak 7281 weight 0 11000E+01 volume 0 11078E+03 ppm1 8 802 ppm2 3 399
(( segid "BRD " and resid 52 and name HN ))
(( segid "BRD " and resid 52 and name H2 ))
ASSI { 7281 } 2 900 2 100 2 100 peak 7291 weight 0 11000E+01 volume 0 28898E+03 ppm1 8 597 ppm2 8 817
(( segid "BRD " and resid 52 and name HN ))
(( segid "BRD " and resid 52 and name H2 ))
ASSI { 7291 } 3 100 2 000 peak 7311 weight 0 11000E+01 volume 0 96479E+02 ppm1 8 794 ppm2 4 770
(( segid "BRD " and resid 52 and name HN ))
(( segid "BRD " and resid 52 and name H2 ))
ASSI { 7301 } 3 200 2 700 2 200 peak 7301 weight 0 11000E+01 volume 0 89979E+02 ppm1 8 594 ppm2 1 393
(( segid "BRD " and resid 52 and name HN ))
(( segid "BRD " and resid 52 and name H2 ))
ASSI { 7311 } 2 600 1 700 1 700 peak 7331 weight 0 11000E+01 volume 0 55160E+03 ppm1 8 599 ppm2 2 454

```

























```

(( segid "BTD " and resid 10 and name H21 ))
(( segid "BTD " and resid 8 and name H2 ))
ASSI { 1553 } 5 200 0 300 peak 1553 weight 0 11000E+01 volume 0 99597E+01 ppm1 8 845 ppm2 4 288 0 759
(( segid "BTD " and resid 7 and name H1 ))
(( segid "BTD " and resid 8 and name H2 ))
ASSI { 1554 } 4 900 0 600 peak 1554 weight 0 11000E+01 volume 0 14339E+02 ppm1 8 924 ppm2 4 381 3 811
(( segid "BTD " and resid 32 and name H1 ))
(( segid "BTD " and resid 32 and name H2 ))
ASSI { 1563 } 3 800 0 700 peak 1563 weight 0 11000E+01 volume 0 61547E+02 ppm1 11 082 ppm2 2 467 3 143
(( segid "BTD " and resid 67 and name H1 ))
(( segid "BTD " and resid 67 and name H2 ))
ASSI { 1564 } 4 600 0 900 peak 1564 weight 0 11000E+01 volume 0 20420E+02 ppm1 8 832 ppm2 4 471 3 077 ppm2 4 143
(( segid "BTD " and resid 89 and name H1 ))
(( segid "BTD " and resid 89 and name H2 ))
ASSI { 1565 } 3 600 0 900 peak 1565 weight 0 11000E+01 volume 0 39185E+02 ppm1 3 669 ppm2 4 556 3 395
(( segid "BTD " and resid 91 and name H1 ))
(( segid "BTD " and resid 91 and name H2 ))
ASSI { 1566 } 3 400 0 900 peak 1566 weight 0 11000E+01 volume 0 32058E+02 ppm1 3 473 ppm2 4 556 3 815 ppm2 4 721
(( segid "BTD " and resid 93 and name H1 ))
(( segid "BTD " and resid 93 and name H2 ))
ASSI { 1567 } 3 400 0 900 peak 1567 weight 0 11000E+01 volume 0 45818E+03 ppm1 4 654 ppm2 3 872 4 654 ppm2 3 872
(( segid "BTD " and resid 94 and name H1 ))
(( segid "BTD " and resid 94 and name H2 ))
ASSI { 1568 } 2 900 0 200 peak 1568 weight 0 11000E+01 volume 0 54263E+02 ppm1 3 473 ppm2 4 411 4 654 ppm2 2 669
(( segid "BTD " and resid 95 and name H1 ))
(( segid "BTD " and resid 95 and name H2 ))
ASSI { 1569 } 3 400 0 900 peak 1569 weight 0 11000E+01 volume 0 54091E+02 ppm1 5 642 ppm2 2 792 3 669 ppm2 5 143
(( segid "BTD " and resid 96 and name H1 ))
(( segid "BTD " and resid 96 and name H2 ))
ASSI { 1570 } 3 400 0 900 peak 1570 weight 0 11000E+01 volume 0 57563E+02 ppm1 3 127 ppm2 7 505 3 522 ppm2 5 143
(( segid "BTD " and resid 97 and name H1 ))
(( segid "BTD " and resid 97 and name H2 ))
ASSI { 1571 } 3 400 0 900 peak 1571 weight 0 11000E+01 volume 0 10961E+03 ppm1 2 733 ppm2 7 506 3 522 ppm2 4 989
(( segid "BTD " and resid 98 and name H1 ))
(( segid "BTD " and resid 98 and name H2 ))
ASSI { 1572 } 3 400 0 900 peak 1572 weight 0 11000E+01 volume 0 23522E+03 ppm1 4 459 ppm2 3 361 4 459 ppm2 3 361
(( segid "BTD " and resid 99 and name H1 ))
(( segid "BTD " and resid 99 and name H2 ))
ASSI { 1573 } 3 400 0 900 peak 1573 weight 0 11000E+01 volume 0 28555E+03 ppm1 3 619 ppm2 4 444 3 619 ppm2 4 444
(( segid "BTD " and resid 100 and name H1 ))
(( segid "BTD " and resid 100 and name H2 ))
ASSI { 1574 } 2 100 0 100 peak 1574 weight 0 11000E+01 volume 0 17185E+03 ppm1 4 409 ppm2 3 998 4 409 ppm2 3 998
(( segid "BTD " and resid 101 and name H1 ))
(( segid "BTD " and resid 101 and name H2 ))
ASSI { 1575 } 2 100 0 100 peak 1575 weight 0 11000E+01 volume 0 16898E+03 ppm1 4 409 ppm2 3 110 4 409 ppm2 3 110
(( segid "BTD " and resid 102 and name H1 ))
(( segid "BTD " and resid 102 and name H2 ))
ASSI { 1576 } 2 100 0 100 peak 1576 weight 0 11000E+01 volume 0 87848E+03 ppm1 2 093 ppm2 4 687 2 093 ppm2 4 687
(( segid "BTD " and resid 103 and name H1 ))
(( segid "BTD " and resid 103 and name H2 ))
ASSI { 1577 } 2 100 0 100 peak 1577 weight 0 11000E+01 volume 0 10742E+04 ppm1 2 190 ppm2 4 440 2 190 ppm2 4 440
(( segid "BTD " and resid 104 and name H1 ))
(( segid "BTD " and resid 104 and name H2 ))
ASSI { 1578 } 2 100 0 100 peak 1578 weight 0 11000E+01 volume 0 84264E+03 ppm1 4 901 ppm2 1 979 4 901 ppm2 1 979
(( segid "BTD " and resid 105 and name H1 ))
(( segid "BTD " and resid 105 and name H2 ))
ASSI { 1579 } 2 100 0 100 peak 1579 weight 0 11000E+01 volume 0 27701E+03 ppm1 5 542 ppm2 4 107 5 542 ppm2 4 107
(( segid "BTD " and resid 106 and name H1 ))
(( segid "BTD " and resid 106 and name H2 ))
ASSI { 1580 } 2 100 0 100 peak 1580 weight 0 11000E+01 volume 0 16922E+03 ppm1 5 542 ppm2 3 146 5 542 ppm2 3 146
(( segid "BTD " and resid 107 and name H1 ))
(( segid "BTD " and resid 107 and name H2 ))
ASSI { 1581 } 2 100 0 100 peak 1581 weight 0 11000E+01 volume 0 10996E+03 ppm1 4 359 ppm2 3 563 4 359 ppm2 3 563
(( segid "BTD " and resid 108 and name H1 ))
(( segid "BTD " and resid 108 and name H2 ))
ASSI { 1582 } 2 100 0 100 peak 1582 weight 0 11000E+01 volume 0 88125E+02 ppm1 4 361 ppm2 3 000 4 361 ppm2 3 000

```







29











35

































(( segid "BTD" * and resid 18 and name HB1 ))	2 400	1 400	0 1000E+01 volume	0 37764E+03 ppm1	2 636 ppm2	2 250	(( segid "BTD" * and resid 76 and name HB1 ))	2 800	2 000	0 1000E+01 volume	0 15977E+03 ppm1	2 092 ppm2	2 564
ASSI [19902]							(( segid "BTD" * and resid 80 and name HB2 ))	OR [19492]					
(( segid "BTD" * and resid 15 and name HB1 ))	2 400	1 400	0 1000E+01 volume	0 35880E+03 ppm1	2 635 ppm2	2 199	(( segid "BTD" * and resid 80 and name HB2 ))	ASSI [19503]					
ASSI [19902]							(( segid "BTD" * and resid 76 and name HB1 ))	ASSI [19503]					
(( segid "BTD" * and resid 75 and name HB1 ))	3 700	3 400	0 1000E+01 volume	0 29255E+02 ppm1	2 636 ppm2	2 444	(( segid "BTD" * and resid 80 and name HB2 ))	ASSI [19522]					
ASSI [19902]							(( segid "BTD" * and resid 76 and name HB1 ))	ASSI [19522]					
(( segid "BTD" * and resid 37 and name HB1 ))	3 200	2 600	0 1000E+01 volume	0 68763E+02 ppm1	2 847 ppm2	4 289	(( segid "BTD" * and resid 76 and name HB1 ))	ASSI [19552]					
ASSI [19902]							(( segid "BTD" * and resid 80 and name HB2 ))	OR [19552]					
(( segid "BTD" * and resid 36 and name HB1 ))	3 400	1 400	0 1000E+01 volume	0 28659E+02 ppm1	2 842 ppm2	5 444	(( segid "BTD" * and resid 76 and name HB1 ))	ASSI [19572]					
ASSI [19902]							(( segid "BTD" * and resid 80 and name HB2 ))	ASSI [19572]					
(( segid "BTD" * and resid 31 and name HB1 ))	3 100	2 400	0 1000E+01 volume	0 98899E+02 ppm1	5 001 ppm2	7 893	(( segid "BTD" * and resid 76 and name HB1 ))	ASSI [19612]					
ASSI [19902]							(( segid "BTD" * and resid 80 and name HB2 ))	ASSI [19612]					
(( segid "BTD" * and resid 30 and name HB1 ))	2 800	2 000	0 1000E+01 volume	0 15610E+03 ppm1	5 001 ppm2	5 428	(( segid "BTD" * and resid 76 and name HB1 ))	ASSI [19642]					
ASSI [19172]							(( segid "BTD" * and resid 80 and name HB2 ))	ASSI [19642]					
(( segid "BTD" * and resid 43 and name HB1 ))	2 900	2 100	0 1000E+01 volume	0 12258E+03 ppm1	5 544 ppm2	2 735	(( segid "BTD" * and resid 76 and name HB1 ))	ASSI [19692]					
ASSI [19172]							(( segid "BTD" * and resid 80 and name HB2 ))	ASSI [19692]					
(( segid "BTD" * and resid 44 and name HB1 ))	3 100	2 400	0 1000E+01 volume	0 97301E+02 ppm1	5 544 ppm2	2 646	(( segid "BTD" * and resid 76 and name HB1 ))	ASSI [19702]					
ASSI [19172]							(( segid "BTD" * and resid 80 and name HB2 ))	ASSI [19702]					
(( segid "BTD" * and resid 43 and name HB1 ))	3 500	3 100	0 1000E+01 volume	0 42319E+02 ppm1	5 544 ppm2	0 799	(( segid "BTD" * and resid 76 and name HB1 ))	ASSI [19722]					
ASSI [19212]							(( segid "BTD" * and resid 80 and name HB2 ))	ASSI [19722]					
(( segid "BTD" * and resid 42 and name HB1 ))	3 200	1 400	0 1000E+01 volume	0 35275E+02 ppm1	1 697 ppm2	2 898	(( segid "BTD" * and resid 76 and name HB1 ))	ASSI [19752]					
ASSI [19242]							(( segid "BTD" * and resid 80 and name HB2 ))	ASSI [19752]					
(( segid "BTD" * and resid 50 and name HB1 ))	5 400	5 400	0 1000E+01 volume	0 31463E+02 ppm1	1 697 ppm2	1 148	(( segid "BTD" * and resid 76 and name HB1 ))	ASSI [19802]					
ASSI [19242]							(( segid "BTD" * and resid 80 and name HB2 ))	ASSI [19802]					
(( segid "BTD" * and resid 39 and name HB1 ))	3 300	2 700	0 1000E+01 volume	0 60793E+02 ppm1	1 697 ppm2	2 032	(( segid "BTD" * and resid 76 and name HB1 ))	ASSI [19812]					
ASSI [19242]							(( segid "BTD" * and resid 80 and name HB2 ))	ASSI [19812]					
(( segid "BTD" * and resid 38 and name HB1 ))	2 200	2 200	0 1000E+01 volume	0 12013E+03 ppm1	1 697 ppm2	1 790	(( segid "BTD" * and resid 76 and name HB1 ))	ASSI [19842]					
ASSI [19312]							(( segid "BTD" * and resid 80 and name HB2 ))	ASSI [19842]					
(( segid "BTD" * and resid 25 and name HB1 ))	3 200	2 600	0 1000E+01 volume	0 68019E+02 ppm1	2 289 ppm2	1 832	(( segid "BTD" * and resid 76 and name HB1 ))	ASSI [19892]					
ASSI [19312]							(( segid "BTD" * and resid 80 and name HB2 ))	ASSI [19892]					
(( segid "BTD" * and resid 31 and name HB1 ))	3 400	2 900	0 1000E+01 volume	0 48169E+02 ppm1	2 289 ppm2	2 841	(( segid "BTD" * and resid 76 and name HB1 ))	ASSI [19922]					
ASSI [19362]							(( segid "BTD" * and resid 80 and name HB2 ))	ASSI [19922]					
(( segid "BTD" * and resid 31 and name HB1 ))	4 200	4 200	0 1000E+01 volume	0 14651E+02 ppm1	2 289 ppm2	3 134	(( segid "BTD" * and resid 76 and name HB1 ))	ASSI [19952]					
ASSI [19362]							(( segid "BTD" * and resid 80 and name HB2 ))	ASSI [19952]					
(( segid "BTD" * and resid 28 and name HB1 ))	1 400	1 400	0 1000E+01 volume	0 37766E+03 ppm1	2 289 ppm2	3 581	(( segid "BTD" * and resid 76 and name HB1 ))	ASSI [19992]					
ASSI [19372]							(( segid "BTD" * and resid 80 and name HB2 ))	ASSI [19992]					
(( segid "BTD" * and resid 35 and name HB1 ))	2 500	1 600	0 1000E+01 volume	0 30400E+03 ppm1	2 289 ppm2	3 457	(( segid "BTD" * and resid 76 and name HB1 ))	ASSI [19992]					
ASSI [19382]							(( segid "BTD" * and resid 80 and name HB2 ))	ASSI [19992]					
(( segid "BTD" * and resid 31 and name HB1 ))	4 200	4 200	0 1000E+01 volume	0 14651E+02 ppm1	2 289 ppm2	3 134	(( segid "BTD" * and resid 76 and name HB1 ))	ASSI [19992]					
ASSI [19392]							(( segid "BTD" * and resid 80 and name HB2 ))	ASSI [19992]					
(( segid "BTD" * and resid 28 and name HB1 ))	3 100	2 400	0 1000E+01 volume	0 87183E+02 ppm1	2 289 ppm2	4 562	(( segid "BTD" * and resid 76 and name HB1 ))	ASSI [19992]					
ASSI [19392]							(( segid "BTD" * and resid 80 and name HB2 ))	ASSI [19992]					
(( segid "BTD" * and resid 31 and name HB1 ))	3 500	3 100	0 1000E+01 volume	0 44223E+02 ppm1	2 289 ppm2	7 803	(( segid "BTD" * and resid 76 and name HB1 ))	ASSI [19992]					
ASSI [19482]							(( segid "BTD" * and resid 80 and name HB2 ))	ASSI [19992]					
(( segid "BTD" * and resid 76 and name HB1 ))	3 300	2 700	0 1000E+01 volume	0 66817E+02 ppm1	2 092 ppm2	2 662	(( segid "BTD" * and resid 76 and name HB1 ))	ASSI [19992]					
ASSI [19492]							(( segid "BTD" * and resid 80 and name HB2 ))	ASSI [19992]					

(( segid "BrD " and reaid 113 and name HA ))	2.800	2.000	2.000 peak 20712 weight	0.10000E+01 volume	0.17530E+03 ppm1	4.656 ppm2	5.111
(( segid "BrD " and reaid 110 and name HA ))	3.500	3.100	2.000 peak 19912 weight	0.10000E+01 volume	0.44050E+02 ppm1	4.903 ppm2	4.419
ASSI [19922]	(( segid "BrD " and reaid 82 and name HA ))	3.900	2.100 peak 20802 weight	0.10000E+01 volume	0.13726E+03 ppm1	1.648 ppm2	7.026
(( segid "BrD " and reaid 113 and name HA ))	3.100	2.400	2.400 peak 19922 weight	0.10000E+01 volume	0.84488E+02 ppm1	4.903 ppm2	4.663
ASSI [19964]	(( segid "BrD " and reaid 113 and name HA ))	3.300	2.200 peak 19992 weight	0.10000E+01 volume	0.60450E+02 ppm1	1.991 ppm2	2.442
(( segid "BrD " and reaid 115 and name HD1 ))	4.000	1.500 peak 19962 weight	0.10000E+01 volume	0.86636E+03 ppm1	1.994 ppm2	2.059	
ASSI [20032]	(( segid "BrD " and reaid 113 and name HA ))	2.100	2.300 peak 20032 weight	0.10000E+01 volume	0.14101E+03 ppm1	1.993 ppm2	4.008
(( segid "BrD " and reaid 113 and name HA ))	2.500	1.400	1.600 peak 20072 weight	0.10000E+01 volume	0.33610E+03 ppm1	4.409 ppm2	1.319
ASSI [20092]	(( segid "BrD " and reaid 17 and name HB ))	3.200	2.600 peak 20092 weight	0.10000E+01 volume	0.76060E+02 ppm1	4.854 ppm2	1.327
(( segid "BrD " and reaid 115 and name HD1 ))	4.600	4.600	9.000 peak 20212 weight	0.10000E+01 volume	0.88717E+03 ppm1	4.542 ppm2	1.224
ASSI [20292]	(( segid "BrD " and reaid 17 and name HB ))	3.500	3.100 peak 20292 weight	0.10000E+01 volume	0.46134E+02 ppm1	1.747 ppm2	2.619
(( segid "BrD " and reaid 115 and name HD1 ))	2.600	1.700	1.700 peak 20242 weight	0.10000E+01 volume	0.27619E+03 ppm1	1.747 ppm2	2.149
ASSI [20252]	(( segid "BrD " and reaid 17 and name HB ))	3.000	2.200 peak 20252 weight	0.10000E+01 volume	0.10098E+03 ppm1	1.747 ppm2	2.507
(( segid "BrD " and reaid 115 and name HD1 ))	4.600	4.600	9.000 peak 20302 weight	0.10000E+01 volume	0.84610E+02 ppm1	1.747 ppm2	0.904
ASSI [20312]	(( segid "BrD " and reaid 39 and name HD1 ))	3.300	2.200 peak 20312 weight	0.10000E+01 volume	0.67487E+02 ppm1	4.903 ppm2	2.271
(( segid "BrD " and reaid 115 and name HD1 ))	3.400	2.900	2.100 peak 20532 weight	0.10000E+01 volume	0.51387E+02 ppm1	1.845 ppm2	5.062
ASSI [20542]	(( segid "BrD " and reaid 41 and name HD1 ))	1.700	1.700 peak 20542 weight	0.10000E+01 volume	0.28354E+03 ppm1	1.844 ppm2	2.019
ASSI [20582]	(( segid "BrD " and reaid 41 and name HD1 ))	2.900	2.100 peak 20582 weight	0.10000E+01 volume	0.13674E+03 ppm1	4.656 ppm2	4.308
ASSI [20332]	(( segid "BrD " and reaid 41 and name HD1 ))	3.400	2.900 peak 20332 weight	0.10000E+01 volume	0.51387E+02 ppm1	1.845 ppm2	5.062
(( segid "BrD " and reaid 115 and name HD1 ))	3.200	2.600	2.300 peak 20572 weight	0.10000E+01 volume	0.75045E+02 ppm1	1.845 ppm2	2.466
ASSI [20592]	(( segid "BrD " and reaid 112 and name HD2 ))	1.400	1.400 peak 20592 weight	0.10000E+01 volume	0.44748E+03 ppm1	2.832 ppm2	4.585
(( segid "BrD " and reaid 67 and name HA ))	3.400	2.900	2.100 peak 21672 weight	0.10000E+01 volume	0.52778E+02 ppm1	1.056 ppm2	7.706













(( segid "Brd" and resid 36 and name HA ))	0.10000E+01 volume	0.54954E+03 ppm1	2.979 ppm2	5.442	(( segid "Brd" and resid 78 and name HD24 ))	0.10000E+01 volume	0.14234E+03 ppm1	0.662 ppm2	7.637
ASSI [27462]	2.900	2.100 peak 27392 weight			ASSI [28012]	2.100	2.100 peak 28012 weight		
(( segid "Brd" and resid 69 and name HD1 ))	0.10000E+01 volume	0.37151E+03 ppm1	1.547 ppm2	1.091	(( segid "Brd" and resid 78 and name HD24 ))	3.500	2.000 peak 28032 weight	0.10000E+01 volume	0.43609E+02 ppm1
ASSI [27462]	2.400	2.400			ASSI [28032]	3.500	2.000 peak 28032 weight	0.10000E+01 volume	0.43609E+02 ppm1
(( segid "Brd" and resid 25 and name HD1 ))	0.10000E+01 volume	0.30318E+03 ppm1	1.599 ppm2	1.788	(( segid "Brd" and resid 78 and name HD24 ))	3.500	2.000 peak 28062 weight	0.10000E+01 volume	0.46405E+02 ppm1
ASSI [27472]	2.500	1.600 peak 27452 weight			ASSI [28062]	3.500	2.000 peak 28062 weight	0.10000E+01 volume	0.46405E+02 ppm1
(( segid "Brd" and resid 22 and name HD1 ))	0.10000E+01 volume	0.54052E+02 ppm1	1.599 ppm2	3.003	(( segid "Brd" and resid 78 and name HD24 ))	4.700	0.800 peak 28112 weight	0.10000E+01 volume	0.77501E+01 ppm1
ASSI [27492]	2.900	2.100 peak 27472 weight			ASSI [28112]	4.700	0.800 peak 28112 weight	0.10000E+01 volume	0.77501E+01 ppm1
(( segid "Brd" and resid 22 and name HD1 ))	0.10000E+01 volume	0.56630E+03 ppm1	1.599 ppm2	4.809	(( segid "Brd" and resid 78 and name HD24 ))	2.900	2.100 peak 28122 weight	0.10000E+01 volume	0.14435E+03 ppm1
ASSI [27552]	2.300	1.300 peak 27492 weight			ASSI [28122]	2.900	2.100 peak 28122 weight	0.10000E+01 volume	0.14435E+03 ppm1
(( segid "Brd" and resid 25 and name HD1 ))	0.10000E+01 volume	0.54070E+03 ppm1	1.599 ppm2	7.559	(( segid "Brd" and resid 78 and name HD24 ))	2.900	2.100 peak 28132 weight	0.10000E+01 volume	0.642 ppm2
ASSI [27572]	2.300	2.200 peak 27552 weight			ASSI [28132]	2.900	2.100 peak 28132 weight	0.10000E+01 volume	0.642 ppm2
(( segid "Brd" and resid 22 and name HD1 ))	0.10000E+01 volume	0.80641E+02 ppm1	1.645 ppm2	4.990	(( segid "Brd" and resid 78 and name HD24 ))	1.800	1.800 peak 28152 weight	0.10000E+01 volume	0.18524E+03 ppm1
ASSI [27602]	2.600	2.300 peak 27572 weight			ASSI [28202]	1.800	1.800 peak 28152 weight	0.10000E+01 volume	0.18524E+03 ppm1
(( segid "Brd" and resid 22 and name HD1 ))	0.10000E+01 volume	0.62418E+03 ppm1	1.646 ppm2	7.529	(( segid "Brd" and resid 78 and name HD24 ))	2.500	1.600 peak 28202 weight	0.10000E+01 volume	0.29008E+03 ppm1
ASSI [27632]	2.200	1.200 peak 27592 weight			ASSI [28232]	2.500	1.600 peak 28202 weight	0.10000E+01 volume	0.29008E+03 ppm1
(( segid "Brd" and resid 73 and name HD1 ))	0.10000E+01 volume	0.31070E+03 ppm1	1.549 ppm2	5.143	(( segid "Brd" and resid 78 and name HD24 ))	3.300	2.200 peak 28232 weight	0.10000E+01 volume	0.64808E+02 ppm1
ASSI [27662]	2.500	1.600 peak 27632 weight			ASSI [28262]	3.300	2.200 peak 28232 weight	0.10000E+01 volume	0.64808E+02 ppm1
(( segid "Brd" and resid 70 and name HD1 ))	0.10000E+01 volume	0.86052E+02 ppm1	1.549 ppm2	4.362	(( segid "Brd" and resid 78 and name HD24 ))	4.100	1.400 peak 28262 weight	0.10000E+01 volume	0.17777E+02 ppm1
ASSI [27712]	3.100	2.400 peak 27662 weight			ASSI [28272]	4.100	1.400 peak 28262 weight	0.10000E+01 volume	0.17777E+02 ppm1
(( segid "Brd" and resid 73 and name HD1 ))	0.10000E+01 volume	0.17668E+03 ppm1	1.500 ppm2	2.106	(( segid "Brd" and resid 78 and name HD24 ))	3.400	2.900 peak 28272 weight	0.10000E+01 volume	0.54803E+02 ppm1
ASSI [27722]	2.800	2.000 peak 27692 weight			ASSI [28292]	3.400	2.900 peak 28272 weight	0.10000E+01 volume	0.54803E+02 ppm1
(( segid "Brd" and resid 35 and name HD1 ))	0.10000E+01 volume	0.22218E+02 ppm1	1.056 ppm2	1.954	(( segid "Brd" and resid 78 and name HD24 ))	3.500	3.100 peak 28312 weight	0.10000E+01 volume	0.41030E+02 ppm1
ASSI [27842]	3.800	3.400 peak 27822 weight			ASSI [28342]	3.500	3.100 peak 28312 weight	0.10000E+01 volume	0.41030E+02 ppm1
(( segid "Brd" and resid 78 and name HD2 ))	0.10000E+01 volume	0.22274E+01 ppm1	1.254 ppm2	4.509	(( segid "Brd" and resid 78 and name HD24 ))	4.000	4.000 peak 28412 weight	0.10000E+01 volume	0.20241E+02 ppm1
ASSI [27892]	2.500	2.500 peak 27872 weight			ASSI [28412]	4.000	4.000 peak 28412 weight	0.10000E+01 volume	0.20241E+02 ppm1
(( segid "Brd" and resid 78 and name HD2 ))	0.10000E+01 volume	0.30077E+03 ppm1	1.254 ppm2	7.047	(( segid "Brd" and resid 78 and name HD24 ))	2.900	2.100 peak 28412 weight	0.10000E+01 volume	0.13708E+03 ppm1
ASSI [27912]	2.600	1.700 peak 27892 weight			ASSI [28472]	2.900	2.100 peak 28412 weight	0.10000E+01 volume	0.13708E+03 ppm1
(( segid "Brd" and resid 78 and name HD2 ))	0.10000E+01 volume	0.24618E+03 ppm1	0.760 ppm2	7.635	(( segid "Brd" and resid 78 and name HD24 ))	2.600	2.300 peak 28472 weight	0.10000E+01 volume	0.70904E+02 ppm1
ASSI [27992]	2.600	1.700 peak 27972 weight			ASSI [28502]	2.600	2.300 peak 28472 weight	0.10000E+01 volume	0.70904E+02 ppm1
(( segid "Brd" and resid 78 and name HD2 ))	0.10000E+01 volume	0.76319E+03 ppm1	0.662 ppm2	7.528	(( segid "Brd" and resid 78 and name HD24 ))	3.600	3.200 peak 28502 weight	0.10000E+01 volume	0.34448E+02 ppm1
ASSI [27992]	2.200	1.200 peak 27972 weight			ASSI [28532]	3.600	3.200 peak 28502 weight	0.10000E+01 volume	0.34448E+02 ppm1
(( segid "Brd" and resid 78 and name HD2 ))	0.10000E+01 volume	0.83417E+03 ppm1	0.662 ppm2	7.031	(( segid "Brd" and resid 78 and name HD24 ))	2.500	1.600 peak 28532 weight	0.10000E+01 volume	0.32948E+03 ppm1
ASSI [27992]	2.100	1.100 peak 27972 weight			ASSI [28562]	2.500	1.600 peak 28532 weight	0.10000E+01 volume	0.32948E+03 ppm1

















Table 3

## Ambiguous NOE-derived Inter-proton Distance Restraints

ASSI { 311 } ( ( segid "BrD" and resid 32 and name HN ) ) ( ( segid "BrD" and resid 35 and name HN ) ) ( ( segid "BrD" and resid 311 weight 0.10000E+01 volume 0.35575E+02 ppm1 OR { 311 } ( ( segid "BrD" and resid 32 and name HN ) ) ( ( segid "BrD" and resid 28 and name H2 ) ) ASSI { 781 } ( ( segid "BrD" and resid 106 and name HN ) ) ( ( segid "BrD" and resid 105 and name HN ) ) ( ( segid "BrD" and resid 105 and name HN ) ) OR { 391 } ( ( segid "BrD" and resid 106 and name HN ) ) ( ( segid "BrD" and resid 104 and name HN ) ) OR { 391 } ( ( segid "BrD" and resid 106 and name HN ) ) ( ( segid "BrD" and resid 107 and name HN ) ) ASSI { 731 } ( ( segid "BrD" and resid 100 and name HN ) ) ( ( segid "BrD" and resid 77 and name HN ) ) ( ( segid "BrD" and resid 77 and name HN ) ) OR { 731 } ( ( segid "BrD" and resid 100 and name HN ) ) ( ( segid "BrD" and resid 100 and name HN ) ) ASSI { 781 } ( ( segid "BrD" and resid 96 and name HN ) ) ( ( segid "BrD" and resid 97 and name HN ) ) ( ( segid "BrD" and resid 97 and name HN ) ) OR { 781 } ( ( segid "BrD" and resid 96 and name HN ) ) ( ( segid "BrD" and resid 97 and name HN ) ) ASSI { 1051 } ( ( segid "BrD" and resid 76 and name HN ) ) ( ( segid "BrD" and resid 72 and name HN ) ) ( ( segid "BrD" and resid 72 and name HN ) ) OR { 1051 } ( ( segid "BrD" and resid 76 and name HN ) ) ( ( segid "BrD" and resid 76 and name HN ) ) ASSI { 1101 } ( ( segid "BrD" and resid 75 and name HN ) ) ( ( segid "BrD" and resid 75 and name HN ) ) ( ( segid "BrD" and resid 75 and name HN ) ) OR { 1101 } ( ( segid "BrD" and resid 75 and name HN ) ) ( ( segid "BrD" and resid 75 and name HN ) ) ASSI { 1551 } ( ( segid "BrD" and resid 21 and name HN ) ) ( ( segid "BrD" and resid 24 and name HN ) ) ( ( segid "BrD" and resid 24 and name HN ) ) OR { 1551 } ( ( segid "BrD" and resid 21 and name HN ) ) ( ( segid "BrD" and resid 21 and name HN ) ) ASSI { 1781 } ( ( segid "BrD" and resid 26 and name HN ) ) ( ( segid "BrD" and resid 26 and name HN ) ) ( ( segid "BrD" and resid 26 and name HN ) ) OR { 1781 } ( ( segid "BrD" and resid 26 and name HN ) ) ( ( segid "BrD" and resid 26 and name HN ) )	3 400 2 900 2 100 peak 1781 weight 0.10000E+01 volume 0.12549E+03 ppm1 ( ( segid "BrD" and resid 26 and name HN ) ) ( ( segid "BrD" and resid 56 and name HN ) ) ASSI { 2721 } ( ( segid "BrD" and resid 104 and name HN ) ) ( ( segid "BrD" and resid 104 and name HN ) ) ( ( segid "BrD" and resid 104 and name HN ) ) OR { 2721 } ( ( segid "BrD" and resid 104 and name HN ) ) ( ( segid "BrD" and resid 104 and name HN ) ) ASSI { 3261 } ( ( segid "BrD" and resid 16 and name HN ) ) ( ( segid "BrD" and resid 13 and name HN ) ) ( ( segid "BrD" and resid 13 and name HN ) ) OR { 3261 } ( ( segid "BrD" and resid 16 and name HN ) ) ( ( segid "BrD" and resid 16 and name HN ) ) ASSI { 3361 } ( ( segid "BrD" and resid 13 and name HN ) ) ( ( segid "BrD" and resid 13 and name HN ) ) ( ( segid "BrD" and resid 13 and name HN ) ) OR { 3361 } ( ( segid "BrD" and resid 13 and name HN ) ) ( ( segid "BrD" and resid 13 and name HN ) ) ASSI { 3371 } ( ( segid "BrD" and resid 14 and name HN ) ) ( ( segid "BrD" and resid 14 and name HN ) ) ( ( segid "BrD" and resid 14 and name HN ) ) OR { 3371 } ( ( segid "BrD" and resid 14 and name HN ) ) ( ( segid "BrD" and resid 14 and name HN ) ) ASSI { 3401 } ( ( segid "BrD" and resid 13 and name HN ) ) ( ( segid "BrD" and resid 13 and name HN ) ) ( ( segid "BrD" and resid 13 and name HN ) ) OR { 3401 } ( ( segid "BrD" and resid 13 and name HN ) ) ( ( segid "BrD" and resid 13 and name HN ) ) ASSI { 3471 } ( ( segid "BrD" and resid 102 and name HN ) ) ( ( segid "BrD" and resid 102 and name HN ) ) ( ( segid "BrD" and resid 102 and name HN ) ) OR { 3471 } ( ( segid "BrD" and resid 102 and name HN ) ) ( ( segid "BrD" and resid 102 and name HN ) ) ASSI { 3841 } ( ( segid "BrD" and resid 100 and name HN ) ) ( ( segid "BrD" and resid 100 and name HN ) ) ( ( segid "BrD" and resid 100 and name HN ) ) OR { 3841 } ( ( segid "BrD" and resid 100 and name HN ) ) ( ( segid "BrD" and resid 100 and name HN ) ) ASSI { 3911 } ( ( segid "BrD" and resid 40 and name HN ) ) ( ( segid "BrD" and resid 41 and name HN ) ) ( ( segid "BrD" and resid 41 and name HN ) ) OR { 3911 } ( ( segid "BrD" and resid 40 and name HN ) ) ( ( segid "BrD" and resid 40 and name HN ) ) ASSI { 4131 } ( ( segid "BrD" and resid 59 and name HN ) ) ( ( segid "BrD" and resid 59 and name HN ) ) ( ( segid "BrD" and resid 59 and name HN ) ) OR { 4131 } ( ( segid "BrD" and resid 59 and name HN ) ) ( ( segid "BrD" and resid 59 and name HN ) ) ASSI { 4661 } ( ( segid "BrD" and resid 111 and name HN ) ) ( ( segid "BrD" and resid 111 and name HN ) ) ( ( segid "BrD" and resid 111 and name HN ) ) OR { 4661 } ( ( segid "BrD" and resid 111 and name HN ) ) ( ( segid "BrD" and resid 111 and name HN ) ) ASSI { 6861 } ( ( segid "BrD" and resid 21 and name HN ) ) ( ( segid "BrD" and resid 24 and name HN ) ) ( ( segid "BrD" and resid 24 and name HN ) ) OR { 6861 } ( ( segid "BrD" and resid 21 and name HN ) ) ( ( segid "BrD" and resid 21 and name HN ) )	9 136 ppm2 2 672 7 763 ppm2 3 671 8 794 ppm2 2 954 8 802 ppm2 3 103 8 810 ppm2 4 488 9 156 ppm2 4 934 8 661 ppm2 4 900 8 496 ppm2 2 969 9 133 ppm2 2 465 8 168 ppm2 2 204 8 574 ppm2 3 072
--	--	--





[illegible]





[illegible]





[illegible]

9









[illegible]

[illegible]

















```

( segid "BrD " and resid 31 and name HB4 )
ASSI { 29702 }
(( segid "BrD " and resid 59 and name HB4 ))
(( segid "BrD " and resid 56 and name HB2 ))
(( segid "BrD " and resid 3400 1 800 peak 29032 weight 0.10000E+01 volume 0 29631E+02 ppm1
OR { 29032 }
(( segid "BrD " and resid 59 and name HB4 ))
(( segid "BrD " and resid 22 and name HB1 ))
OR { 29032 }
(( segid "BrD " and resid 59 and name HB4 ))
(( segid "BrD " and resid 21 and name HB1 ))
ASSI { 29702 }
(( segid "BrD " and resid 59 and name HB4 ))
(( segid "BrD " and resid 22 and name HB1 ))
(( segid "BrD " and resid 4200 4 200 1 300 peak 29072 weight 0 10000E+01 volume 0 14615E+02 ppm1
OR { 29072 }
(( segid "BrD " and resid 59 and name HB4 ))
(( segid "BrD " and resid 56 and name HB2 ))
OR { 29072 }
(( segid "BrD " and resid 59 and name HB4 ))
(( segid "BrD " and resid 59 and name HB4 ))
OR { 29072 }
(( segid "BrD " and resid 59 and name HB4 ))
(( segid "BrD " and resid 59 and name HB4 ))
ASSI { 29152 }
(( segid "BrD " and resid 59 and name HB4 ))
(( segid "BrD " and resid 73 and name HB2 ))
(( segid "BrD " and resid 3 600 3 200 1 900 peak 29152 weight 0 10000E+01 volume 0 3825E+02 ppm1
OR { 29152 }
(( segid "BrD " and resid 59 and name HB4 ))
(( segid "BrD " and resid 59 and name HB4 ))
(( segid "BrD " and resid 59 and name HB4 ))
OR { 29152 }
(( segid "BrD " and resid 59 and name HB4 ))
(( segid "BrD " and resid 59 and name HB4 ))
ASSI { 29192 }
(( segid "BrD " and resid 54 and name HB4 ))
(( segid "BrD " and resid 60 and name HB1 ))
(( segid "BrD " and resid 3200 2 600 2 300 peak 29192 weight 0 10000E+01 volume 0 7636E+02 ppm1
OR { 29192 }
(( segid "BrD " and resid 54 and name HB4 ))
(( segid "BrD " and resid 77 and name HB1 ))
(( segid "BrD " and resid 54 and name HB4 ))
(( segid "BrD " and resid 54 and name HB4 ))
(( segid "BrD " and resid 35 3 700 3 400 1 800 peak 29202 weight 0 10000E+01 volume 0 30281E+02 ppm1
OR { 29202 }
(( segid "BrD " and resid 54 and name HB4 ))
(( segid "BrD " and resid 37 and name HB1 ))
ASSI { 29202 }
(( segid "BrD " and resid 54 and name HB4 ))
(( segid "BrD " and resid 54 and name HB4 ))
(( segid "BrD " and resid 54 and name HB4 ))
OR { 29252 }
(( segid "BrD " and resid 54 and name HB4 ))
(( segid "BrD " and resid 61 and name HB2 ))
OR { 29252 }
(( segid "BrD " and resid 54 and name HB4 ))
(( segid "BrD " and resid 54 and name HB4 ))
ASSI { 29262 }
(( segid "BrD " and resid 54 and name HB4 ))
(( segid "BrD " and resid 37 and name HB1 ))
(( segid "BrD " and resid 1 700 1 700 peak 29262 weight 0 10000E+01 volume 0 2310E+03 ppm1
OR { 29262 }
(( segid "BrD " and resid 54 and name HB4 ))
(( segid "BrD " and resid 54 and name HB4 ))
ASSI { 29502 }
(( segid "BrD " and resid 79 and name HB4 ))
(( segid "BrD " and resid 116 and name HB1 ))
(( segid "BrD " and resid 4 100 4 100 1 400 peak 29502 weight 0 10000E+01 volume 0 16991E+02 ppm1
OR { 29502 }
(( segid "BrD " and resid 79 and name HB4 ))
(( segid "BrD " and resid 103 and name HB2 ))
ASSI { 29602 }
(( segid "BrD " and resid 108 and name HB4 ))
(( segid "BrD " and resid 111 and name HB1 ))
(( segid "BrD " and resid 4200 4 100 1 400 peak 29602 weight 0 10000E+01 volume 0 1647E+02 ppm1
OR { 29602 }
(( segid "BrD " and resid 108 and name HB4 ))
(( segid "BrD " and resid 109 and name HB1 ))
ASSI { 29702 }
(( segid "BrD " and resid 15 and name HB4 ))
(( segid "BrD " and resid 102 and name HB2 ))
(( segid "BrD " and resid 2 000 2 000 peak 29702 weight 0 10000E+01 volume 0 16053E+03 ppm1
OR { 29702 }
(( segid "BrD " and resid 108 and name HB4 ))

```

```

OR { 29702 }
(( segid "BrD " and resid 109 and name HB1 ))
(( segid "BrD " and resid 60 and name HB2 ))
(( segid "BrD " and resid 56 and name HB2 ))
ASSI { 29712 }
(( segid "BrD " and resid 35 and name HA ))
(( segid "BrD " and resid 16 and name HB1 ))
(( segid "BrD " and resid 2 400 2 400 peak 29712 weight 0 10000E+01 volume 0 22638E+02 ppm1
OR { 29712 }
(( segid "BrD " and resid 60 and name HB2 ))
(( segid "BrD " and resid 64 and name HB1 ))
(( segid "BrD " and resid 15 and name HA ))
(( segid "BrD " and resid 14 and name HB2 ))
(( segid "BrD " and resid 3 800 1 600 peak 29722 weight 0 10000E+01 volume 0 22638E+02 ppm1
OR { 29722 }
(( segid "BrD " and resid 15 and name HA ))
(( segid "BrD " and resid 65 and name HB2 ))
ASSI { 29732 }
(( segid "BrD " and resid 30 and name HA ))
(( segid "BrD " and resid 98 and name HA ))
(( segid "BrD " and resid 1 700 1 700 peak 29732 weight 0 10000E+01 volume 0 26724E+03 ppm1
OR { 29732 }
(( segid "BrD " and resid 30 and name HA ))
(( segid "BrD " and resid 29 and name HA ))
ASSI { 30592 }
(( segid "BrD " and resid 109 and name HB1 ))
(( segid "BrD " and resid 21 and name HB1 ))
(( segid "BrD " and resid 3 400 1 800 peak 30592 weight 0 10000E+01 volume 0 31262E+02 ppm1
OR { 30592 }
(( segid "BrD " and resid 109 and name HB1 ))
(( segid "BrD " and resid 110 and name HB2 ))
ASSI { 30672 }
(( segid "BrD " and resid 109 and name HB2 ))
(( segid "BrD " and resid 110 and name HA ))
(( segid "BrD " and resid 3 200 1 900 peak 30672 weight 0 10000E+01 volume 0 39278E+02 ppm1
OR { 30672 }
(( segid "BrD " and resid 109 and name HB2 ))
(( segid "BrD " and resid 107 and name HA ))
ASSI { 30952 }
(( segid "BrD " and resid 78 and name HA ))
(( segid "BrD " and resid 46 and name HB3 ))
(( segid "BrD " and resid 47 and name HA ))
(( segid "BrD " and resid 3 600 1 700 peak 30952 weight 0 10000E+01 volume 0 28677E+02 ppm1
OR { 30952 }
(( segid "BrD " and resid 26 and name HA ))
(( segid "BrD " and resid 56 and name HB2 ))
ASSI { 83 }
(( segid "BrD " and resid 46 and name HB3 ))
(( segid "BrD " and resid 47 and name HA ))
(( segid "BrD " and resid 38 and name HB ))
(( segid "BrD " and resid 2 600 2 300 peak 533 weight 0 10000E+01 volume 0 34245E+02 ppm1
OR { 533 }
(( segid "BrD " and resid 47 and name HB4 ))
(( segid "BrD " and resid 50 and name HB ))
ASSI { 533 }
(( segid "BrD " and resid 74 and name HB4 ))
(( segid "BrD " and resid 71 and name HA ))
(( segid "BrD " and resid 3 100 2 400 2 400 peak 643 weight 0 10000E+01 volume 0 35398E+02 ppm1
OR { 643 }
(( segid "BrD " and resid 74 and name HB4 ))
(( segid "BrD " and resid 15 and name HA ))
OR { 643 }
(( segid "BrD " and resid 74 and name HB4 ))
(( segid "BrD " and resid 76 and name HA ))
OR { 643 }
(( segid "BrD " and resid 74 and name HB4 ))
(( segid "BrD " and resid 58 and name HB4 ))
OR { 643 }
(( segid "BrD " and resid 74 and name HB4 ))
(( segid "BrD " and resid 72 and name HA ))
OR { 643 }
(( segid "BrD " and resid 74 and name HB4 ))
(( segid "BrD " and resid 14 and name HB ))
ASSI { 1193 }
(( segid "BrD " and resid 105 and name HB4 ))
(( segid "BrD " and resid 101 and name HB2 ))
(( segid "BrD " and resid 2 400 2 400 peak 1143 weight 0 10000E+01 volume 0 36898E+02 ppm1
OR { 1143 }
(( segid "BrD " and resid 105 and name HB4 ))
(( segid "BrD " and resid 21 and name HB2 ))
ASSI { 1393 }
(( segid "BrD " and resid 86 and name HB4 ))
(( segid "BrD " and resid 96 and name HA ))
(( segid "BrD " and resid 2 200 2 200 peak 1393 weight 0 10000E+01 volume 0 26898E+03 ppm1
OR { 1393 }

```

4 607 ppm2 2 182

4 607 ppm2 1 417

5.445 ppm2 4 810

2 334 ppm2 1 222

2 141 ppm2 4 378

4.509 ppm2 1 260

5 758 ppm2 4 693

7 246 ppm2 1 776

7 012 ppm2 4 622

7 758 ppm2 1 621

7 711 ppm2 4 810

7 617 ppm2	4 431
7 617 ppm2	7 027
7 478 ppm2	4 810
7 618 ppm2	4 387
7 804 ppm2	4 971
6 687 ppm2	1 777
7 266 ppm2	4 424
7 270 ppm2	2 504
7 970 ppm2	0 780
7 970 ppm2	4 007
8 158 ppm2	4 538

784	673	755	212	579	791	440	333	791	683
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----









```

( segid "BrD " and resid 38 and name H021)
ASSI ( 4144) "BrD " and resid 74 and name H01 )
(( segid "BrD " and resid 78 and name H01 ))
OR ( 4144) "BrD " and resid 82 and name H01 ))
ASSI ( 4154) "BrD " and resid 78 and name H01 ))
(( segid "BrD " and resid 74 and name H01 ))
(( segid "BrD " and resid 22 and name H01 ))
OR ( 4154) "BrD " and resid 106 and name H01 ))
ASSI ( 4174) "BrD " and resid 21 and name H01 ))
(( segid "BrD " and resid 96 and name H01 ))
(( segid "BrD " and resid 100 and name H01 ))
OR ( 4174) "BrD " and resid 34 and name H01 ))
ASSI ( 4224) "BrD " and resid 32 and name H01 ))
(( segid "BrD " and resid 50 and name H01 ))
OR ( 4224) "BrD " and resid 95 and name H01 ))
ASSI ( 4234) "BrD " and resid 96 and name H01 ))
(( segid "BrD " and resid 99 and name H01 ))
OR ( 4234) "BrD " and resid 106 and name H01 ))
ASSI ( 4244) "BrD " and resid 106 and name H01 ))
(( segid "BrD " and resid 115 and name H01 ))
OR ( 4244) "BrD " and resid 106 and name H01 ))
ASSI ( 4284) "BrD " and resid 74 and name H01 ))
(( segid "BrD " and resid 78 and name H01 ))
OR ( 4284) "BrD " and resid 106 and name H01 ))
ASSI ( 4344) "BrD " and resid 109 and name H01 ))
(( segid "BrD " and resid 106 and name H01 ))
OR ( 4344) "BrD " and resid 106 and name H01 ))
ASSI ( 4354) "BrD " and resid 74 and name H01 ))
(( segid "BrD " and resid 18 and name H01 ))

```

0.10000E+01 volume 0.77150E+02 ppm1 7.005 ppm2 1.249

0.10000E+01 volume 0.70341E+03 ppm1 7.534 ppm2 1.676

0.10000E+01 volume 0.13533E+03 ppm1 7.689 ppm2 4.927

0.10000E+01 volume 0.12637E+03 ppm1 7.615 ppm2 0.845

0.10000E+01 volume 0.94963E+02 ppm1 7.611 ppm2 2.178

0.10000E+01 volume 0.14612E+03 ppm1 7.616 ppm2 0.763

0.10000E+01 volume 0.18031E+03 ppm1 7.534 ppm2 3.010

0.10000E+01 volume 0.12963E+03 ppm1 7.904 ppm2 5.555

0.10000E+01 volume 0.19182E+03 ppm1 7.539 ppm2 1.089



Table 4

## Hydrogen Bonding Restraints

!Helix Z						
assign (residue 19 and name HN )	(residue 15 and name O )	1.80	0.0	0.40		
assign (residue 19 and name N )	(residue 15 and name O )	2.80	0.30	0.40		
assign (residue 22 and name HN )	(residue 18 and name O )	1.80	0.0	0.40		
assign (residue 22 and name N )	(residue 18 and name O )	2.80	0.30	0.40		
assign (residue 23 and name HN )	(residue 19 and name O )	1.80	0.0	0.40		
assign (residue 23 and name N )	(residue 19 and name O )	2.80	0.30	0.40		
assign (residue 24 and name HN )	(residue 20 and name O )	1.80	0.0	0.40		
assign (residue 24 and name N )	(residue 20 and name O )	2.80	0.30	0.40		
assign (residue 25 and name HN )	(residue 21 and name O )	1.80	0.0	0.40		
assign (residue 25 and name N )	(residue 21 and name O )	2.80	0.30	0.40		
!Helix B						
assign (residue 75 and name HN )	(residue 71 and name O )	1.80	0.0	0.40		
assign (residue 75 and name N )	(residue 71 and name O )	2.80	0.30	0.40		
!assign (residue 77 and name HN )	(residue 73 and name O )	1.80	0.0	0.40		
!assign (residue 77 and name N )	(residue 73 and name O )	2.80	0.30	0.40		
assign (residue 78 and name HN )	(residue 74 and name O )	1.80	0.0	0.40		
assign (residue 78 and name N )	(residue 74 and name O )	2.80	0.30	0.40		
assign (residue 79 and name HN )	(residue 75 and name O )	1.80	0.0	0.40		
assign (residue 79 and name N )	(residue 75 and name O )	2.80	0.30	0.40		
!assign (residue 80 and name HN )	(residue 76 and name O )	1.80	0.0	0.40		
!assign (residue 80 and name N )	(residue 76 and name O )	2.80	0.30	0.40		
assign (residue 81 and name HN )	(residue 77 and name O )	1.80	0.0	0.40		
assign (residue 81 and name N )	(residue 77 and name O )	2.80	0.30	0.40		
assign (residue 82 and name HN )	(residue 78 and name O )	1.80	0.0	0.40		
assign (residue 82 and name N )	(residue 78 and name O )	2.80	0.30	0.40		
!Helix C						
assign (residue 102 and name HN )	(residue 98 and name O )	1.80	0.0	0.40		
assign (residue 102 and name N )	(residue 98 and name O )	2.80	0.30	0.40		
assign (residue 103 and name HN )	(residue 99 and name O )	1.80	0.0	0.40		
assign (residue 103 and name N )	(residue 99 and name O )	2.80	0.30	0.40		
assign (residue 104 and name HN )	(residue 100 and name O )	1.80	0.0	0.40		
assign (residue 104 and name N )	(residue 100 and name O )	2.80	0.30	0.40		
assign (residue 105 and name HN )	(residue 101 and name O )	1.80	0.0	0.40		
assign (residue 105 and name N )	(residue 101 and name O )	2.80	0.30	0.40		

Table 5

# Atomic Structure Coordinates of the Free Form of the P/CAF Bromodomain

REMARK FILENAME="/bioch2/chrs/BROW\_MPRO\_ARIA32/structures/ica/brd187.pdb"  
REMARK initial random number seed 1 142876511  
REMARK  
REMARK overall\_bonds\_angles\_180  
REMARK  
REMARK energy: 157.9231, 9.0626, 73.1523, 0.23, 1819, 36.4277, 0.228429  
REMARK  
REMARK bonds, angles, improper, noe, cdh  
REMARK  
REMARK rms-dev 2.164158E-01, 0.365111, 50.3111, 1.418958E-02, 0.263503  
REMARK  
REMARK noe, cdh  
REMARK violations - 2, 0  
REMARK DATE 29-Nov-98 06:51:33 created by user

REMARK DATE 20-Nov-08									
ATOM	1	CA	GLY	1	27.208	16.825	-6.349	1.00	0.00
ATOM	2	HA	GLY	1	28.043	15.521	-5.365	1.00	0.00
ATOM	3	HA2	GLY	1	28.043	15.521	-5.365	1.00	0.00
ATOM	4	C	GLY	1	27.703	15.432	-6.650	1.00	0.00
ATOM	5	O	GLY	1	28.656	15.263	-7.435	1.00	0.00
ATOM	6	N	GLY	1	28.194	17.285	-7.337	1.00	0.00
ATOM	7	H11	GLY	1	26.586	18.049	-7.923	1.00	0.00
ATOM	8	H12	GLY	1	26.586	17.641	-6.943	1.00	0.00
ATOM	9	H13	GLY	1	25.930	17.641	-6.943	1.00	0.00
ATOM	10	N	SER	2	27.113	14.431	-6.623	1.00	0.00
ATOM	11	HN	SER	2	26.376	14.629	-5.409	1.00	0.00
ATOM	12	CA	SER	2	27.516	13.044	-6.236	1.00	0.00
ATOM	13	HA	SER	2	28.481	13.047	-6.712	1.00	0.00
ATOM	14	HA2	SER	2	27.638	12.329	-4.878	1.00	0.00
ATOM	15	H11	SER	2	27.638	12.329	-4.878	1.00	0.00
ATOM	16	H12	SER	2	27.462	11.273	-5.017	1.00	0.00
ATOM	17	H3	SER	2	27.462	11.273	-5.017	1.00	0.00
ATOM	18	OG	SER	2	28.927	12.508	-4.319	1.00	0.00
ATOM	19	HG	SER	2	29.571	12.014	-4.832	1.00	0.00
ATOM	20	O	SER	2	26.519	12.308	-7.114	1.00	0.00
ATOM	21	C	SER	2	27.020	12.288	-6.935	1.00	0.00
ATOM	22	HN	HIS	3	27.986	11.747	-8.163	1.00	0.00
ATOM	23	CA	HIS	3	26.173	10.955	-9.106	1.00	0.00
ATOM	24	HA	HIS	3	25.157	11.295	-8.967	1.00	0.00
ATOM	25	HA2	HIS	3	26.594	11.222	-10.553	1.00	0.00
ATOM	26	H11	HIS	3	27.616	12.288	-10.723	1.00	0.00
ATOM	27	H12	HIS	3	26.616	12.288	-10.723	1.00	0.00
ATOM	28	CG	HIS	3	26.671	10.612	-11.561	1.00	0.00
ATOM	29	ND1	HIS	3	25.985	10.494	-12.900	1.00	0.00
ATOM	30	ND2	HIS	3	26.820	10.790	-13.300	1.00	0.00
ATOM	31	C	HIS	3	24.433	10.060	-11.420	1.00	0.00
ATOM	32	H12	HIS	3	24.970	9.918	-13.501	1.00	0.00
ATOM	33	CG1	HIS	3	24.961	9.918	-13.501	1.00	0.00
ATOM	34	HE1	HIS	3	24.945	9.698	-14.594	1.00	0.00
ATOM	35	NE2	HIS	3	24.028	9.658	-12.662	1.00	0.00
ATOM	36	HE2	HIS	3	23.171	9.230	-12.868	1.00	0.00
ATOM	37	CG	HIS	3	26.233	9.459	-8.817	1.00	0.00
ATOM	38	O	HIS	3	26.224	8.638	-9.734	1.00	0.00
ATOM	39	N	MET	4	26.304	9.311	-7.534	1.00	0.00
ATOM	40	HN	MET	4	26.314	9.111	-6.850	1.00	0.00
ATOM	41	CA	MET	4	26.364	7.713	-7.128	1.00	0.00
ATOM	42	HA	MET	4	26.164	7.105	-7.998	1.00	0.00
ATOM	43	HA2	MET	4	27.755	7.374	-6.591	1.00	0.00
ATOM	44	CB	MET	4	28.456	8.117	-6.944	1.00	0.00
ATOM	45	HB1	MET	4	27.710	7.400	-5.512	1.00	0.00
ATOM	46	HB2	MET	4	27.710	7.400	-5.512	1.00	0.00
ATOM	47	CG	MET	4	28.284	6.005	-7.024	1.00	0.00
ATOM	48	H01	MET	4	27.433	5.306	-6.972	1.00	0.00
ATOM	49	H02	MET	4	28.606	6.071	-8.043	1.00	0.00
ATOM	50	SD	MET	4	29.597	5.394	-5.988	1.00	0.00
ATOM	51	HE1	MET	4	29.243	3.623	-4.435	1.00	0.00
ATOM	52	HE2	MET	4	29.273	3.039	-6.102	1.00	0.00
ATOM	53	HE3	MET	4	27.843	3.866	-5.483	1.00	0.00
ATOM	54	C	MET	4	25.310	7.413	-6.067	1.00	0.00
ATOM	55	O	MET	4	25.333	7.983	-4.977	1.00	0.00
ATOM	56	N	SER	5	24.386	6.517	-6.399	1.00	0.00
ATOM	57	HN	SER	5	24.422	6.098	-7.283	1.00	0.00
ATOM	58	CA	SER	5	23.321	6.141	-5.476	1.00	0.00
ATOM	59	HA	SER	5	23.780	5.750	-4.580	1.00	0.00
ATOM	60	H11	SER	5	22.479	7.366	-5.111	1.00	0.00
ATOM	61	H12	SER	5	22.188	7.882	-6.014	1.00	0.00
ATOM	62	H01	SER	5	23.062	8.039	-4.490	1.00	0.00
ATOM	63	OG	SER	5	21.309	6.950	-4.405	1.00	0.00
ATOM	64	HG	SER	5	21.060	7.693	-3.800	1.00	0.00
ATOM	65	C	SER	5	22.431	5.062	-6.084	1.00	0.00
ATOM	66	O	SER	5	21.471	5.362	-6.194	1.00	0.00

[illegible]



ATOM	819	CD	ARG	51	-2.433	4.849	-11.240	1.00	0.00	ATOM	914	CG	LVS	57	-5.542	-10.395	9.520	1.00	0.00	ATOM	1007	C	ARG	62	4.700	-11.542	-0.564	1.00	0.00
ATOM	821	HD2	ARG	51	-1.296	5.010	-11.240	1.00	0.00	ATOM	916	HN	LEU	57	-6.160	-11.395	1.629	1.00	0.00	ATOM	1008	N	ARG	62	5.791	-11.240	-0.564	1.00	0.00
ATOM	822	HD2	ARG	51	-1.566	5.493	-11.244	1.00	0.00	ATOM	918	HN	LEU	57	-7.622	-10.954	2.517	1.00	0.00	ATOM	1009	N	LEU	63	4.700	-10.366	-0.053	1.00	0.00
ATOM	823	HD2	ARG	51	-1.275	5.184	-12.385	1.00	0.00	ATOM	919	CD	LVS	57	-6.049	-11.720	3.735	1.00	0.00	ATOM	1010	HN	LEU	63	3.882	-9.835	-0.768	1.00	0.00
ATOM	824	CD	ARG	51	-2.632	5.236	-12.212	1.00	0.00	ATOM	920	HD1	LVS	57	-5.944	-11.047	4.573	1.00	0.00	ATOM	1011	CA	LEU	63	5.792	-9.853	0.768	1.00	0.00
ATOM	825	HN	LEU	51	-1.562	4.953	-13.904	1.00	0.00	ATOM	921	HD1	LVS	57	-5.917	-12.837	3.498	1.00	0.00	ATOM	1012	HA	LEU	63	6.834	-9.663	0.137	1.00	0.00
ATOM	826	HN1	ARG	51	-1.231	4.984	-13.846	1.00	0.00	ATOM	922	HN2	LVS	57	-7.967	-12.640	6.335	1.00	0.00	ATOM	1014	HB2	LEU	63	5.208	-7.805	-0.805	1.00	0.00
ATOM	827	HN1	ARG	51	-0.936	4.716	-13.161	1.00	0.00	ATOM	923	HB2	LVS	57	-6.621	-13.771	3.759	1.00	0.00	ATOM	1015	HB2	LEU	63	4.460	-8.716	1.983	1.00	0.00
ATOM	828	HN1	ARG	51	-3.559	5.539	-14.026	1.00	0.00	ATOM	924	HN2	LVS	57	-7.223	-12.935	5.985	1.00	0.00	ATOM	1016	CG	LEU	63	6.413	-7.984	2.436	1.00	0.00
ATOM	829	HN2	ARG	51	-3.559	5.539	-14.026	1.00	0.00	ATOM	925	HN2	LVS	57	-6.343	-12.609	6.084	1.00	0.00	ATOM	1017	HG	LEU	63	7.141	-8.747	2.663	1.00	0.00
ATOM	830	HD2	ARG	51	-4.617	5.746	-14.430	1.00	0.00	ATOM	926	HN2	LVS	57	-6.343	-12.609	6.084	1.00	0.00	ATOM	1018	HN1	LEU	63	6.413	-7.984	2.436	1.00	0.00
ATOM	831	C	ARG	51	-3.694	6.354	-11.056	1.00	0.00	ATOM	927	O	LVS	57	-3.855	-10.344	1.617	1.00	0.00	ATOM	1019	HD1	LEU	63	7.575	-7.097	0.877	1.00	0.00
ATOM	832	O	ARG	51	-3.479	-0.182	-9.901	1.00	0.00	ATOM	928	N	LVS	57	-3.197	-11.146	1.895	1.00	0.00	ATOM	1021	HD1	LEU	63	7.937	-6.487	2.492	1.00	0.00
ATOM	833	N	SER	52	-3.370	-0.652	-9.134	1.00	0.00	ATOM	929	N	THR	58	-4.014	-9.714	0.370	1.00	0.00	ATOM	1022	CD2	LEU	63	5.781	-7.580	3.738	1.00	0.00
ATOM	834	N	THR	58	-3.370	-0.652	-9.134	1.00	0.00	ATOM	930	HN	THR	58	-4.351	-8.913	0.201	1.00	0.00	ATOM	1023	HD21	LEU	63	4.709	-7.899	3.726	1.00	0.00
ATOM	835	HA	SER	52	-3.766	-2.101	-10.603	1.00	0.00	ATOM	931	HN	THR	58	-4.351	-8.913	0.201	1.00	0.00	ATOM	1024	HN2	LEU	63	5.285	-6.507	4.848	1.00	0.00
ATOM	836	HA	SER	52	-3.766	-2.101	-10.603																						





ATOM	1666	HD13	ILE	101	-6.076	5.842	11.709	1.00	0.00	7.384	11.534	1.00	0.00	106	1759	CE2	PHE	106	1853	NZ	LVS	111	7.830	11.169	2.702	1.00	0.00	
ATOM	1667	C	ILE	101	-2.511	6.457	8.232	1.00	0.00	7.360	11.560	1.00	0.00	107	1760	CE2	PHE	107	1854	NZ	LVS	112	7.360	11.560	3.543	1.00	0.00	
ATOM	1668	C	ILE	101	-1.335	6.807	8.232	1.00	0.00	7.331	11.531	1.00	0.00	108	1761	CE2	PHE	108	1855	H23	LVS	111	8.121	11.554	2.084	1.00	0.00	
ATOM	1669	HA	LEU	102	-2.876	5.343	7.590	1.00	0.00	5.125	5.780	1.050	1.00	0.00	109	1762	HZ	PHE	109	1856	H23	LVS	111	8.121	11.554	2.084	1.00	0.00
ATOM	1670	HA	LEU	102	-2.876	5.343	7.590	1.00	0.00	5.125	5.780	1.050	1.00	0.00	110	1763	C	PHE	110	1857	C	LVS	111	12.448	7.070	4.425	1.00	0.00
ATOM	1671	CA	LEU	102	-1.886	4.457	6.792	1.00	0.00	5.368	5.033	5.140	1.00	0.00	111	1764	C	PHE	111	1858	C	LVS	111	12.448	7.070	4.425	1.00	0.00
ATOM	1672	HA	LEU	102	-1.350	3.933	6.762	1.00	0.00	5.610	5.033	5.140	1.00	0.00	112	1765	HN	PHE	112	1859	C	LVS	111	12.448	7.070	4.425	1.00	0.00
ATOM	1673	CA	LEU	102	-1.350	3.933	6.762	1.00	0.00	5.610	5.033	5.140	1.00	0.00	113	1766	HN	PHE	113	1860	C	LVS	111	12.448	7.070	4.425	1.00	0.00
ATOM	1674	CB	LEU	102	-2.569	3.416	6.082	1.00	0.00	5.075	6.216	2.885	1.00	0.00	114	1767	HA	PHE	114	1861	CB	GLU	112	14.297	6.205	6.161	1.00	0.00
ATOM	1675	HA	LEU	102	-3.630	3.443	6.283	1.00	0.00	5.662	6.216	2.885	1.00	0.00	115	1768	HA	PHE	115	1862	HA	GLU	112	14.297	6.205	6.161	1.00	0.00
ATOM	1676	CB	LEU	102	-3.630	3.443	6.283	1.00	0.00	5.662	6.216	2.885	1.00	0.00	116	1769	HA	PHE	116	1863	HA	GLU	112	14.297	6.205	6.161	1.00	0.00
ATOM	1677	HD	LEU	102	-2.423	1.362	5.420	1.00	0.00	2.918	7.770	1.838	1.00	0.00	117	1770	HD	PHE	117	1864	HD	GLU	112	13.126	4.819	7.296	1.00	0.00
ATOM	1678	HD	LEU	102	-2.423	1.362	5.420	1.00	0.00	2.918	7.770	1.838	1.00	0.00	118	1771	HD	PHE	118	1865	HD	GLU	112	13.126	4.819	7.296	1.00	0.00
ATOM	1679	HD11	LEU	102	-1.943	1.395	8.316	1.00	0.00	3.038	6.475	1.348	1.00	0.00	119	1772	CD1	PHE	119	1866	HD2	GLU	112	12.440	5.320	6.289	1.00	0.00
ATOM	1680	HD11	LEU	102	-1.943	1.395	8.316	1.00	0.00	3.038	6.475	1.348	1.00	0.00	120	1773	CD1	PHE	120	1867	HD2	GLU	112	12.440	5.320	6.289	1.00	0.00
ATOM	1681	HD11	LEU	102	-2.958	0.338	7.332	1.00	0.00	2.098	8.860	1.160	1.00	0.00	121	1774	CD1	PHE	121	1868	HD2	GLU	112	12.440	5.320	6.289	1.00	0.00
ATOM	1682	CD2	LEU	102	-0.559	1.963	6.290	1.00	0.00																			





Table 6

# Atomic Structure Coordinates of the P/CAF Bromodomain/Acetyl-Histamine Complex

FILENAME=~/bin/chr2/chr2/complex_xplor_aria13/structures/alc/complex_83	REMARK
DATE=20140905 15:05:00	REMARK
overall_bonds_angle Improper_vdw_noe cdh1REMARK energies 154 107, 9 #62626,	REMARK
72 1621, 0, 22.2303, 36 0151, 0 204924REMARK	REMARK
Bonds_angle Improper_vdw_noe cdh1REMARK energies 154 107, 9 #62626,	REMARK
72 1621, 0, 22.2303, 36 0151, 0 204924REMARK	REMARK
noe, cdh1REMARK violations, 2, 0REMARK	REMARK
Dec-98 15:11:47 created by user ATOM	REMARK
2 661 -4.968 1.00 0.00	1 CA ACE 200
2 661 -4.968 1.00 0.00	2 HB2 ACE 200
2 304 -5.428 1.00 0.00	3 HB2 ACE 200
3 572 -5.067 1.00 0.00	4 HB2 ACE 200
2 236 -3.376 1.00 0.00	5 O ACE 200
2 238 -3.032 1.00 0.00	6 O ACE 200
3 967 -2.463 1.00 0.00	7 N H1M 201
3 967 -2.463 1.00 0.00	8 N H1M 201
3 827 -1.417 1.00 0.00	9 CA H1M 201
3 827 -1.417 1.00 0.00	10 H1M 201
3 589 -0.620 1.00 0.00	11 H1M 201
3 025 -1.526 1.00 0.00	12 CB H1M 201
5 100 -1.055 1.00 0.00	13 H1M 201
5 940 -1.107 1.00 0.00	14 H1M 201
5 940 -1.107 1.00 0.00	15 CB H1M 201
5 329 -2.010 1.00 0.00	16 ND1 H1M 201
6 259 -1.993 1.00 0.00	17 ND1 H1M 201
4 956 -1.057 1.00 0.00	18 CD2 H1M 201
4 756 -3.204 1.00 0.00	19 CD2 H1M 201
4 116 -3.779 1.00 0.00	20 CD2 H1M 201
6 732 -2.908 1.00 0.00	21 ND2 H1M 201
5 221 -3.592 1.00 0.00	22 ND2 H1M 201
6 667 -0.366 1.00 0.00	23 CA GLY 1
16.138 -0.871 1.00 0.00	24 HA2 GLY 1
15.945 -0.255 1.00 0.00	25 HA2 GLY 1
17.144 -1.411 1.00 0.00	26 C GLY 1
17.144 -1.411 1.00 0.00	27 N GLY 1
17.800 -1.198 1.00 0.00	28 HT1 GLY 1
17.963 -1.018 1.00 0.00	29 HT2 GLY 1
17.584 -2.207 1.00 0.00	30 HT3 GLY 1
16.682 -0.970 1.00 0.00	31 N SER 2
16.682 -0.970 1.00 0.00	32 N SER 2
16.682 -0.970 1.00 0.00	33 N SER 2
17.043 -3.411 1.00 0.00	34 CA SER 2
17.546 -3.401 1.00 0.00	35 HA SER 2
17.546 -3.401 1.00 0.00	36 CB SER 2
18.588 -4.762 1.00 0.00	37 HB1 SER 2
17.414 -4.374 1.00 0.00	38 HB2 SER 2
18.429 -2.660 1.00 0.00	39 HB3 SER 2
18.429 -2.660 1.00 0.00	40 HB3 SER 2
18.429 -2.660 1.00 0.00	41 C SER 2
18.429 -2.660 1.00 0.00	42 O SER 2
18.429 -2.660 1.00 0.00	43 N HIS 3
18.429 -2.660 1.00 0.00	44 HN HIS 3
18.429 -2.660 1.00 0.00	45 HN HIS 3
18.429 -2.660 1.00 0.00	46 HN HIS 3
18.429 -2.660 1.00 0.00	47 CB HIS 3
18.429 -2.660 1.00 0.00	48 HB1 HIS 3
18.429 -2.660 1.00 0.00	49 HB2 HIS 3
18.429 -2.660 1.00 0.00	50 CB HIS 3
18.429 -2.660 1.00 0.00	51 ND1 HIS 3
18.429 -2.660 1.00 0.00	52 ND2 HIS 3
18.429 -2.660 1.00 0.00	53 CD2 HIS 3
18.429 -2.660 1.00 0.00	54 HB2 HIS 3
18.429 -2.660 1.00 0.00	55 CB HIS 3
18.429 -2.660 1.00 0.00	56 HB1 HIS 3
18.429 -2.660 1.00 0.00	57 HB2 HIS 3
18.429 -2.660 1.00 0.00	58 HB3 HIS 3
18.429 -2.660 1.00 0.00	59 C HIS 3
18.429 -2.660 1.00 0.00	60 O HIS 3
18.429 -2.660 1.00 0.00	61 N MET 4
18.429 -2.660 1.00 0.00	62 HN MET 4
18.429 -2.660 1.00 0.00	63 HN MET 4
18.429 -2.660 1.00 0.00	64 HA MET 4















4.534	5.217	1.00	0.00	BxD ATOM	1945	CD1	ILE	116	4.662
0.635	6.273	1.00	0.00	BxD ATOM	1946	HD1	ILE	116	1.000
1.238	7.147	1.00	0.00	BxD ATOM	1947	HD13	ILE	116	5.246
0.086	6.812	1.00	0.00	BxD ATOM	1949	C	ILE	116	5.356
1.751	8.053	1.00	0.00	BxD ATOM	1950	O	ILE	116	6.802
0.544	8.154	1.00	0.00	BxD ATOM	1951	N	ASP	117	7.024
2.453	9.014	1.00	0.00	BxD ATOM	1952	EN	ASP	117	6.222
2.453	9.014	1.00	0.00	BxD ATOM	1952	EN	ASP	117	5.905
1.862	10.275	1.00	0.00	BxD ATOM	1954	HA	ASP	117	5.197
2.585	10.798	1.00	0.00	BxD ATOM	1955	CB	ASP	117	4.967
0.609	10.013	1.00	0.00	BxD ATOM	1956	HB1	ASP	117	5.564
0.287	10.220	1.00	0.00	BxD ATOM	1957	HB2	ASP	117	4.664
0.596	8.976	1.00	0.00	BxD ATOM	1958	HB1	ASP	117	1.000
1.488	10.787	1.00	0.00	BxD ATOM	1958	CD1	ASP	117	2.497
0.421	11.641	1.00	0.00	BxD ATOM	1961	C	ASP	117	1.572
1.506	11.154	1.00	0.00	BxD ATOM	1962	O	ASP	117	7.006
0.892	12.209	1.00	0.00	BxD ATOM	1963	N	LVS	118	6.846
1.897	10.725	1.00	0.00	BxD ATOM	1964	HN	LVS	118	8.206
1.897	10.725	1.00	0.00	BxD ATOM	1964	HN	LVS	118	8.282
1.613	11.868	1.00	0.00	BxD ATOM	1965	HA	LVS	118	10.248
2.053	10.968	1.00	0.00	BxD ATOM	1967	CB	LVS	118	9.315
2.230	12.889	1.00	0.00	BxD ATOM	1968	HB1	LVS	118	8.636
1.624	13.480	1.00	0.00	BxD ATOM	1969	HB2	LVS	118	10.292
2.199	13.349	1.00	0.00	BxD ATOM	1970	CD	LVS	118	9.818
3.719	12.296	1.00	0.00	BxD ATOM	1971	HB1	LVS	118	8.595
3.719	12.296	1.00	0.00	BxD ATOM	1972	HB2	LVS	118	8.595
3.946	13.911	1.00	0.00	BxD ATOM	1973	CD	LVS	118	9.857
4.608	12.327	1.00	0.00	BxD ATOM	1974	HD1	LVS	118	10.477
4.972	13.133	1.00	0.00	BxD ATOM	1975	HB2	LVS	118	10.468
4.075	11.613	1.00	0.00	BxD ATOM	1976	CB	LVS	118	9.208
5.425	10.902	1.00	0.00	BxD ATOM	1977	HB1	LVS	118	8.595
6.386	12.369	1.00	0.00	BxD ATOM	1979	HB2	LVS	118	8.287
6.652	10.943	1.00	0.00	BxD ATOM	1980	HB1	LVS	118	10.212
6.458	11.316	1.00	0.00	BxD ATOM	1981	HB2	LVS	118	11.163
7.658	11.096	1.00	0.00	BxD ATOM	1982	HB2	LVS	118	9.989
0.130	11.891	1.00	0.00	BxD ATOM	1983	HB1	LVS	118	10.206
0.641	10.802	1.00	0.00	BxD ATOM	1984	OT1	LVS	118	9.047
0.302	12.459	1.00	0.00	BxD END	1985	OT2	LVS	118	10.454

## Structure-based sequence homology alignment of bromodomains

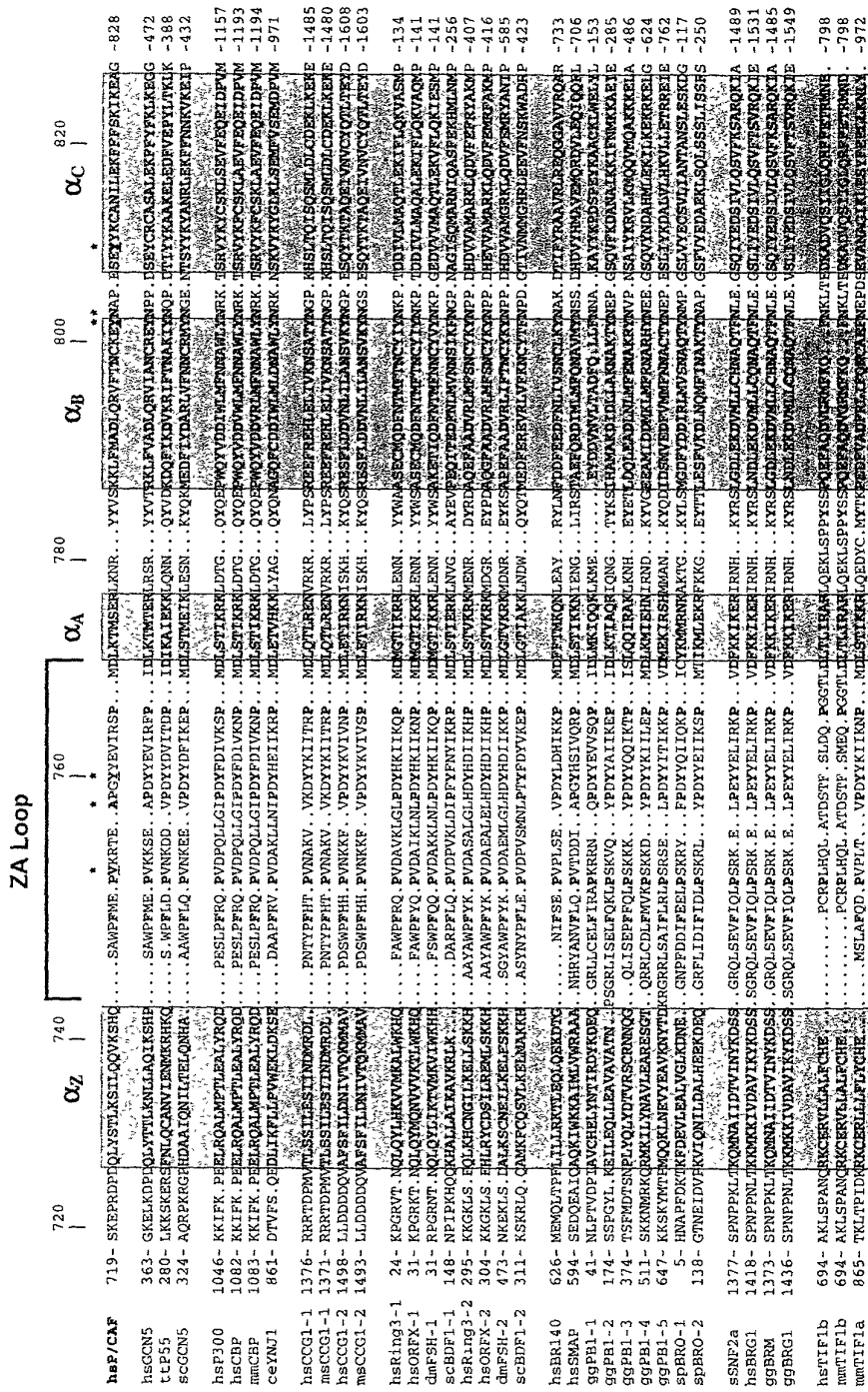


Figure 1

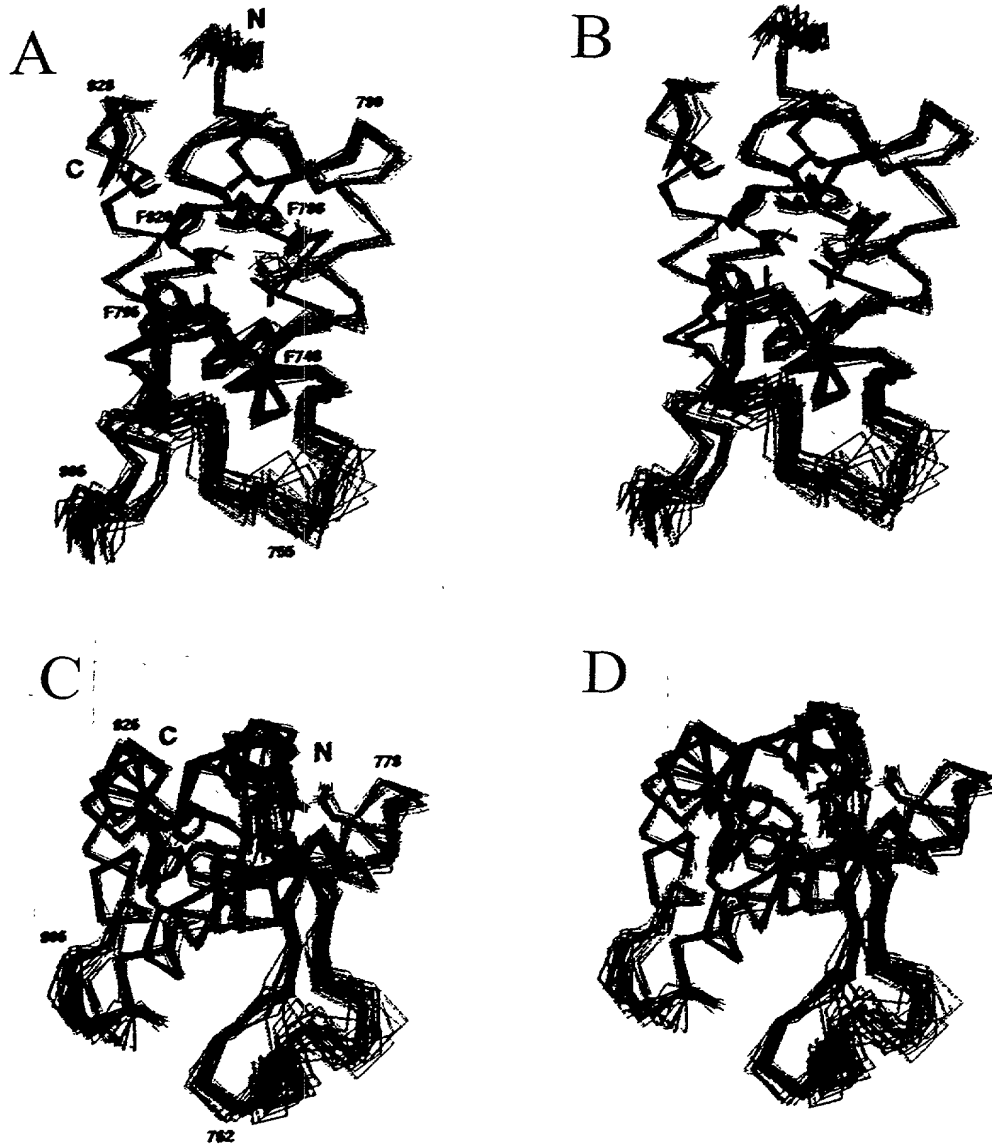


Figure 2A-2D

**Three-Dimensional Structure of the P/CAF Bromodomain**

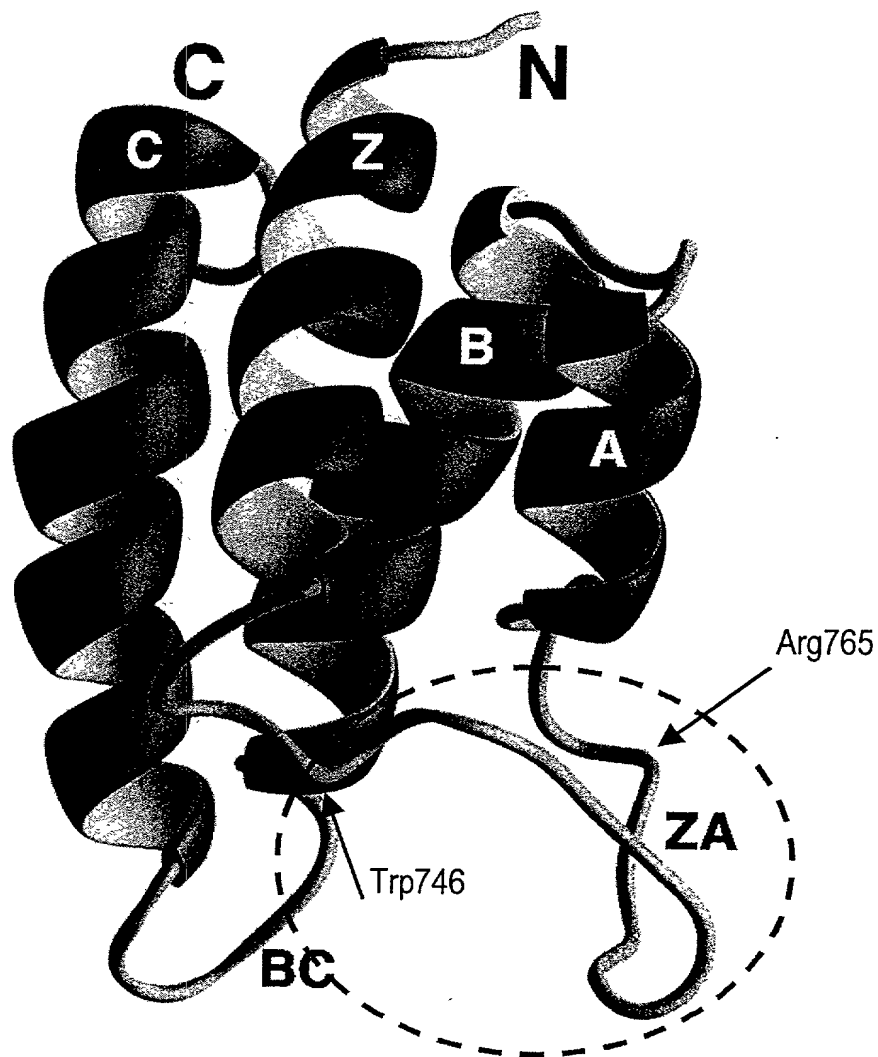


Figure 2E

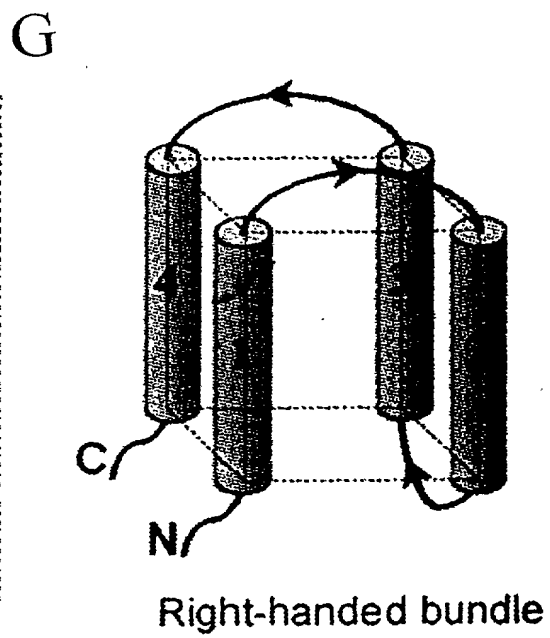
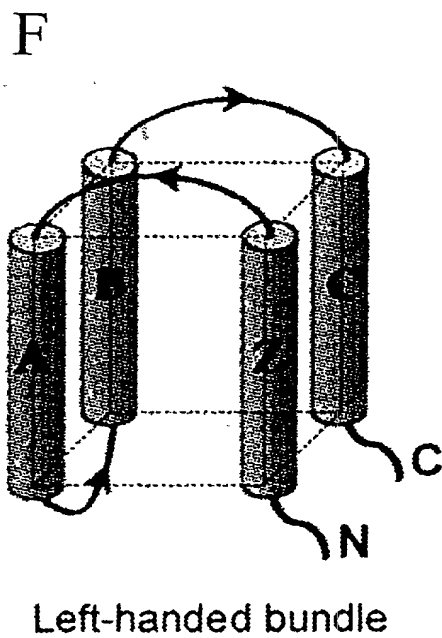


Figure 2F-2G

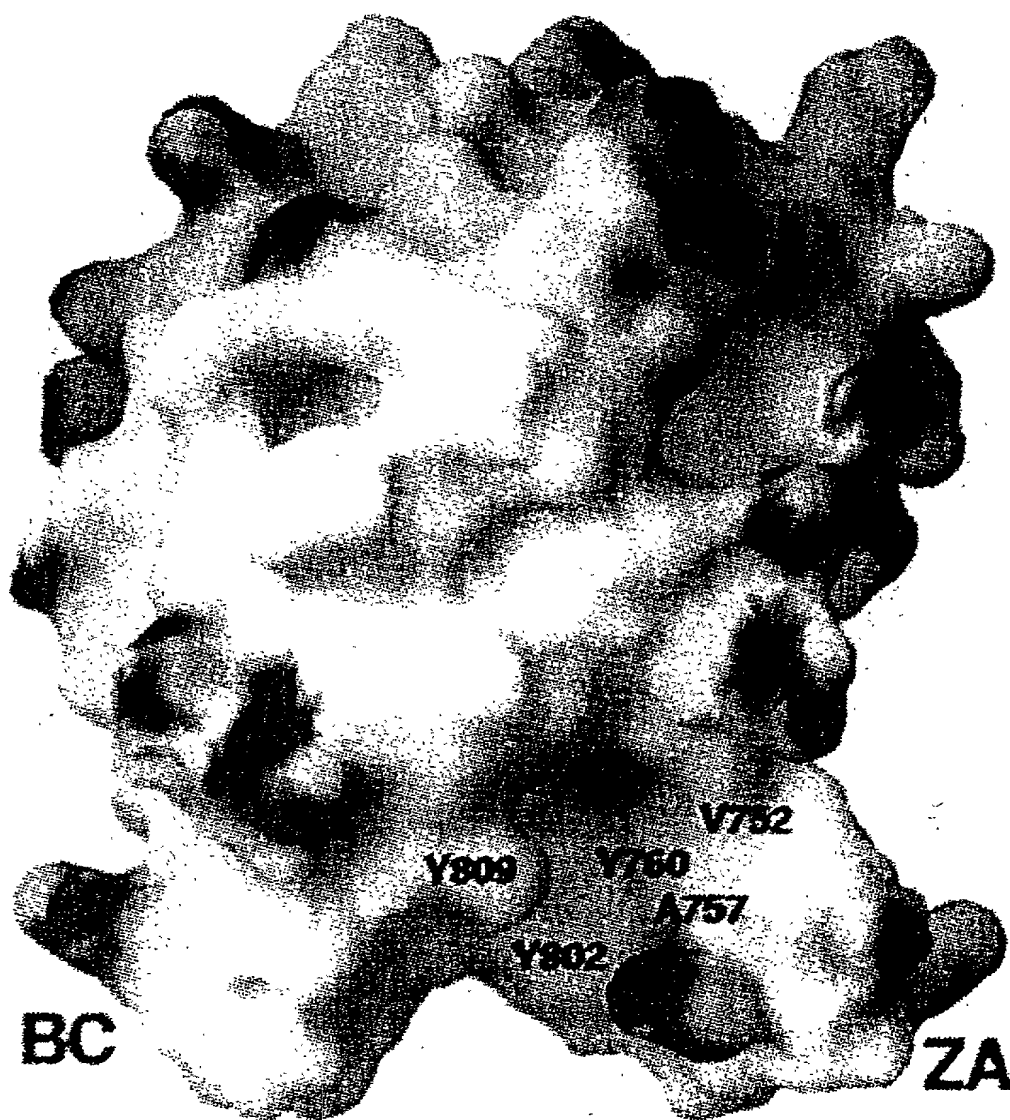


Figure 2H

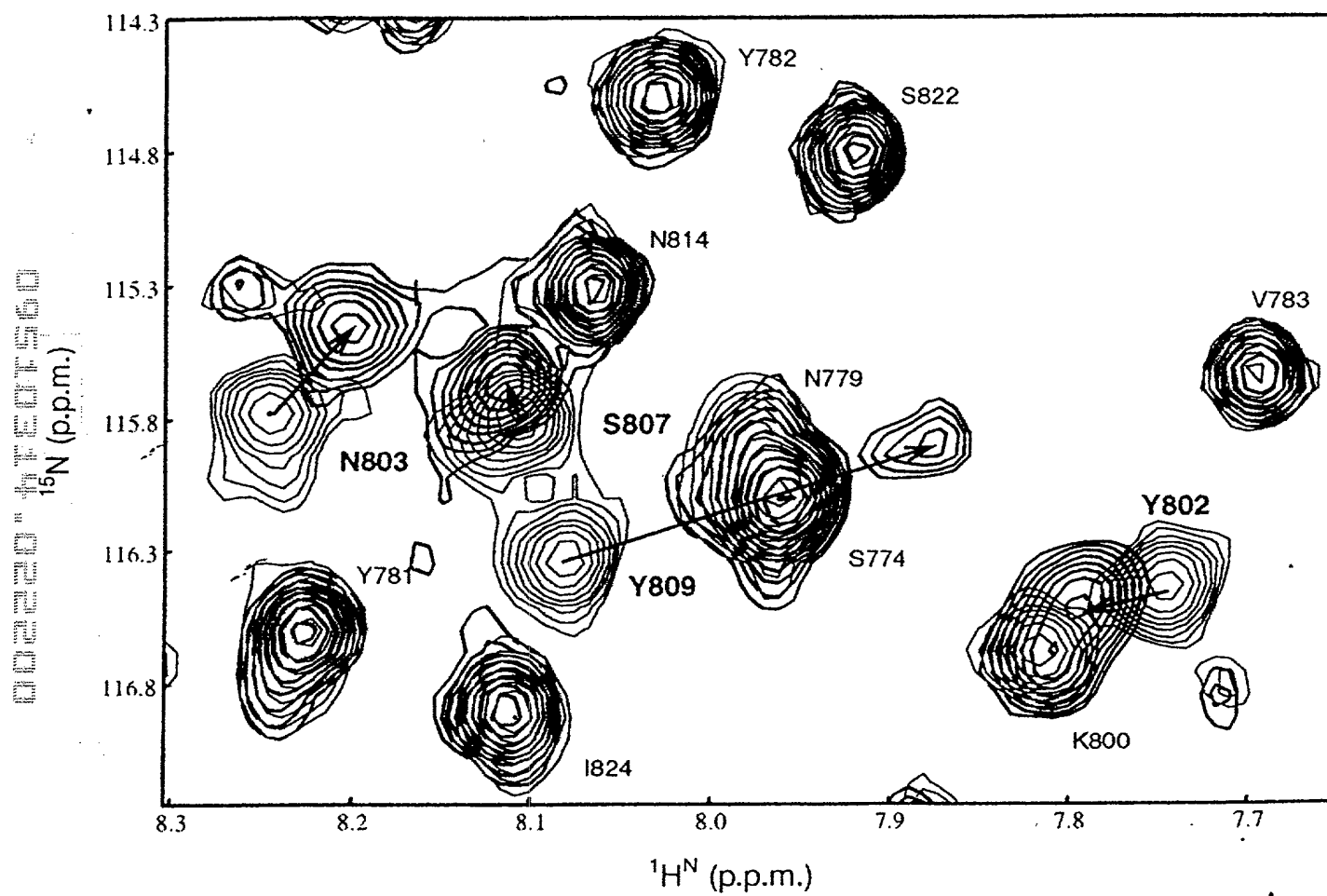


Figure 3A

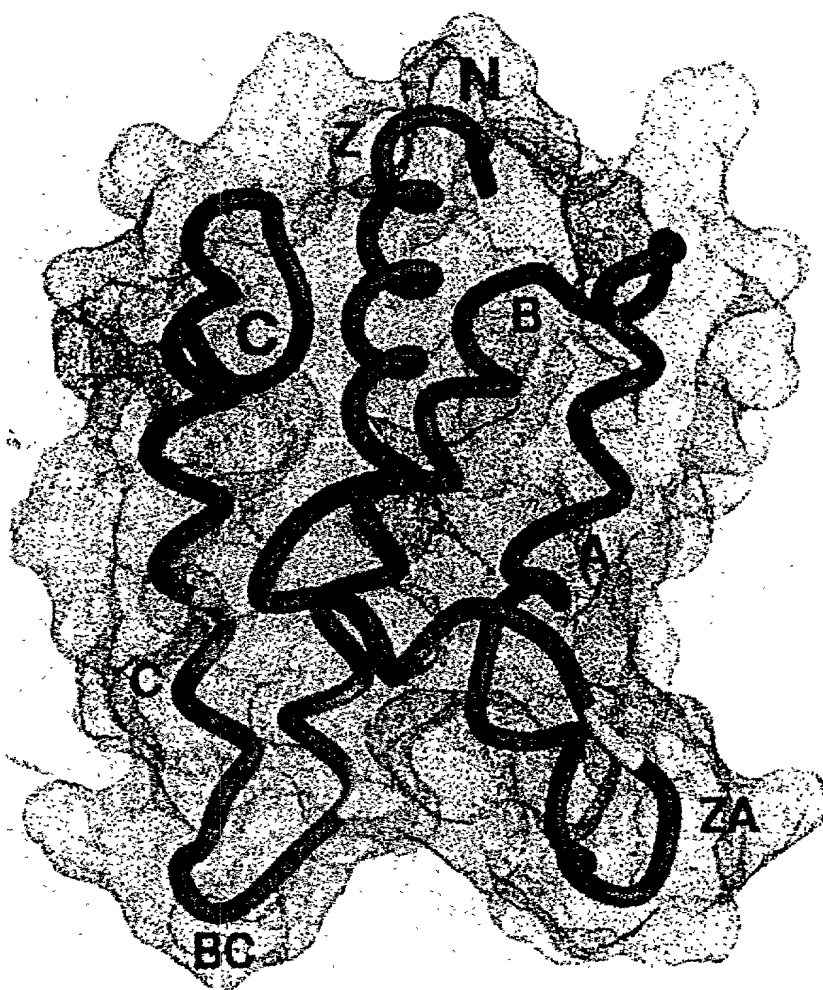
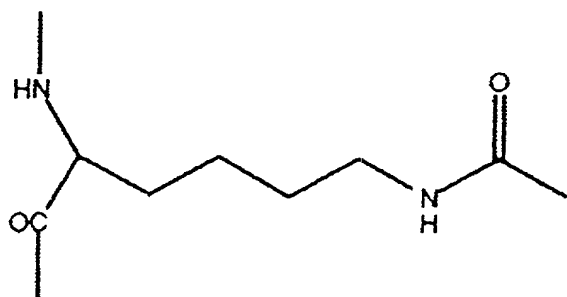
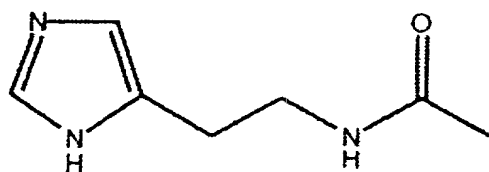


Figure 3B

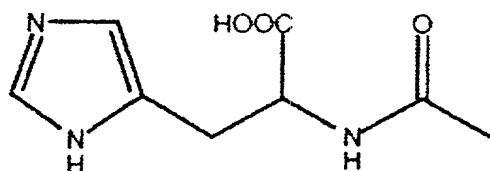




N $\epsilon$ -acetyl-lysine



N $\omega$ -acetyl-histamine



N $\alpha$ -acetyl-histidine

Figure 3C

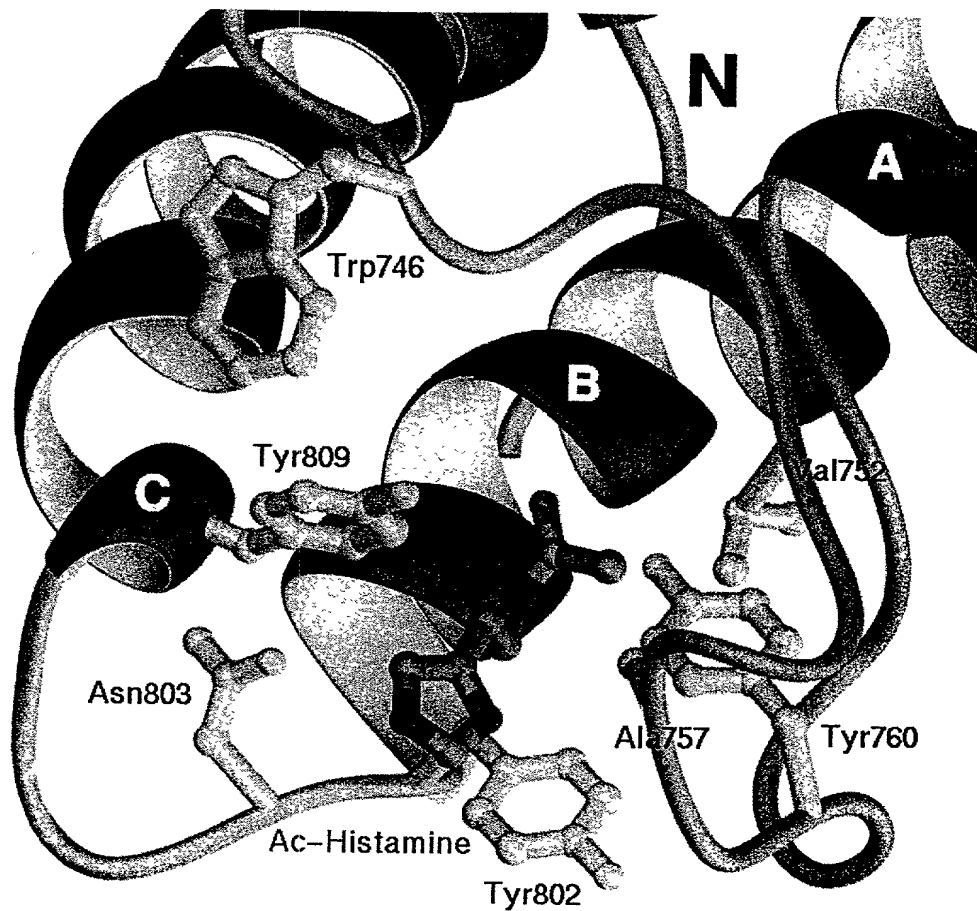


Figure 4

**DECLARATION AND POWER OF ATTORNEY FOR PATENT APPLICATION**

As below named inventors, we hereby declare that:

Our residence, post office address and citizenship are as stated below under our names.

We believe that we are the original, first and joint inventors of the subject matter which is claimed and for which a patent is sought on the invention entitled

**METHODS OF IDENTIFYING MODULATORS OF BROMODOMAINS**

the Specification of which

☒ is attached hereto  
☐ was filed on \_\_\_\_\_  
as Application Serial No. \_\_\_\_\_  
and was amended on \_\_\_\_\_ (if applicable).

We hereby state that we have reviewed and understand the contents of the above-identified Specification, including the claims, as amended by any amendment referred to above.

We acknowledge the duty to disclose information which is material to the examination of this application in accordance with Title 37, Code of Federal Regulations, 1.56(a).

We hereby claim foreign priority benefits under Title 35, United States Code, §119 of any provisional application filed in the United States in accordance with 35 U.S.C. §1.119(e), or any application for patent that has been converted to a Provisional Application within one (1) year of its filing date, or any foreign application(s) for patent or inventor's certificate listed below and have also identified below any foreign application for patent or inventor's certificate having a filing date before that of the application on which priority is claimed.

**PRIOR FILED APPLICATION(S)**

<u>APPLICATION</u> <u>NUMBER</u>	<u>COUNTRY</u>	<u>(DAY/MONTH/YEAR FILED)</u>	<u>PRIORITY</u> <u>CLAIMED</u>
-------------------------------------	----------------	-------------------------------	-----------------------------------

We hereby claim the benefit under Title 35, United States Code, §120 of any United States application listed below, and, insofar as the subject matter of each of the claims of this application is not disclosed in any prior United States application in the manner provided by the first paragraph of Title 35, United States Code, §112, I acknowledge the duty to disclose material information as defined in Title 37, Code of Federal Regulations, §1.56(a), which occurred between the filing date of the prior application and the national or PCT international filing date of this application:

APPLICATION NO. \_\_\_\_\_ FILING DATE (DAY/MONTH/YEAR) STATUS - PATENTED, PENDING, ABANDONED \_\_\_\_\_

We hereby appoint as our attorneys or agents the following persons: Stefan J. Klauber (Attorney, Registration No. 22,604); David A. Jackson (Attorney, Registration No. 26,742); Donald J. Cox, Jr. (Attorney, Registration No. 37,804); Michael D. Davis (Attorney, Registration No. 39,161); Allan H. Fried (Attorney, Registration No. 31,253); Christine E. Dietzel (Agent, Registration No. 37,309); and Michael A. Yamin (Agent, Registration No. P44,414), said attorneys or agents with full power of substitution and revocation to prosecute this application and transact all business in the Patent and Trademark Office connected therewith.

Please address all correspondence regarding this application to:

DAVID A. JACKSON, ESQ.  
KLAUBER & JACKSON  
411 HACKENSACK AVENUE  
HACKENSACK, NEW JERSEY 07601

Direct all telephone calls to David A. Jackson at (201) 487-5800.

We hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further, that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

FULL NAME OF FIRST JOINT INVENTOR: MING-MING ZHOU  
COUNTRY OF CITIZENSHIP: People's Republic of China  
FULL RESIDENCE ADDRESS: 35 Richmond Drive  
Greenwich, CT 06870  
FULL POST OFFICE ADDRESS: 35 Richmond Drive  
Greenwich, CT 06870

SIGNATURE OF INVENTOR \_\_\_\_\_

DATE \_\_\_\_\_



# SEQUENCE LISTING

<110> Zhou, Ming-Ming  
Aggarnal, Aneel K

<120> METHODS OF IDENTIFYING MODULATORS OF BROMODOMAINS

<130> 2459-1-003

<140> UNASSIGNED

<141> 2000-02-22

<160> 44

<170> PatentIn Ver. 2.0

<210> 1

<211> 3014

<212> DNA

<213> Homo sapiens

<400> 1

```

ggggccgcgt cgacgcgga aagaggccgt ggggggcctc ccagcgctgg cagacaccgt 60
gaggctggca gccgccggca cgcacaccta gtccgcagtc ccgaggaaca tgtccgcagc 120
cagggcgcg agcagagtcc cgggcaggag aaccaaggga gggcggtgtc tgtggcgggc 180
gcggcagcgg cagcggagcc gctagtcccc tccctcctgg gggagcagct gccgccgctg 240
ccgccgccgc caccaccatc agcgcgcggg gcccggccag agcagagccg gcgagcggcg 300
cgctaggggg agggcggggg cggggagggg ggtgggcgaa gggggcgggg gggcggtggg 360
ggaggggtct gctctcccga ctaccagagc ccgagggaga ccctggcggc ggcggcgggc 420
cctgacactc ggcgcctcct gccgtgtctc ggggcggcat gtccgaggct ggcggggccg 480
ggccggggcg ctgcggggca ggagccgggg cagggggccg gcccgggcg ctgccccgc 540
agcctgcggc gcttcgcgcc gcgccccgc agggctcccc ctgcgcgct gccgccgggg 600
gctcggggcg ctgcggtccg gcgacggcag tggctgcagc gggcacggcc gaaggaccgg 660
gaggcgggtg ctgcggcccga atcgccgtga agaaagcgca actacgctcc gctccgcggg 720
ccaagaaact ggagaaactc ggagtgtact ccgcctgcaa ggccgaggag tcttgtaaat 780
gtaatggctg gaaaaaccct aaccctcac ccactcccc cagagccgac ctgcagcaaa 840
taattgtcag tctaacagaa tcctgtcgga gttgtagcca tgccctagct gctcatgttt 900
cccacctgga gaatgtgtca gaggaagaaa tgaacagact cctgggaata gtattggatg 960
tggaatatct ctttacctgt gtccacaagg aagaagatgc agataccaaa caagtttatt 1020
tctatctatt taagctcttg agaaagtcta ttttacaagg aggaaaacct gtggttgaag 1080
gctctttgga aaagaaaccc ccatttgaaa aacctagcat tgaacagggt gtgaataact 1140
ttgtgcagta caaathtagt cacctgccag caaaagaaag gcaaacaata gttgagttgg 1200
caaaaatgtt cctaaaccgc atcaactatt ggcattctgga ggcaccatct caacgaagac 1260
tgcatctcc caatgatgat atttctggat acaaagagaa ctacacaagg tggctgtgtt 1320
actgaacgt gccacagttc tgcgacagtc tacctcggtc cgaaaccaca cagggtgttt 1380
ggagaacatt gcttcgctcg gtcttactg ttatgaggcg acaactcctg gaacaagcaa 1440
gacaggaaaa agataaactg cctcttgaaa aacgaactct aatcctcact catttccaa 1500
aatttctgtc catgctagaa gaagaagtat atagtcaaaa ctctcccatc tgggatcagg 1560

```

```

attttctctc agcctcttcc agaaccagcc agctaggcat ccaaacagtt atcaatccac 1620
ctcctgtggc tgggacaatt tcatacaatt caacctcatc ttcccttgag cagccaaacg 1680
cagggagcag cagtcctgcc tgcaaagcct cttctggact tgaggcaaac ccaggagaaa 1740
agaggaaaat gactgattct catgttctgg aggaggccaa gaaaccccga gttatggggg 1800
atattccgat ggaattaatc aacgaggtta tgtctaccat cacggaccct gcagcaatgc 1860
ttggaccaga gaccaatttt ctgtcagcac actcggccag ggatgaggcg gcaaggttgg 1920
aagagcgcag ggggtgtaatt gaatttcacg tggttggcaa ttccctcaac cagaaaccaa 1980
acaagaagat cctgatgtgg ctggttggcc tacagaacgt tttctcccac cagctgcccc 2040
gaatgccaaa agaatacatc acacggctcg tctttgaccc gaaacacaaa acccttgctt 2100
taattaaaga tggccgtggt attggttggt tctgtttccg tatgttccca tctcaaggat 2160
tcacagagat tgtcttctgt gctgtaacct caaatgagca agtcaaggcg tatggaacac 2220
acctgatgaa tcatttgaaa gaatatcaca taaagcatga catcctgaac ttcttcacat 2280
atgcagatga atatgcaatt ggatacttta agaaacaggg tttctccaaa gaaattaaaa 2340
tacctaaaac caaatatggt ggctatatca aggattatga aggagccact ttaatgggat 2400
gtgagctaaa tccacggatc ccgtacacag aattttctgt catcattaaa aagcagaagg 2460
agataattaa aaaactgatt gaaagaaaac aggcacaaat tcgaaaagtt taccctggac 2520
tttcatgttt taaagatgga gttcgacaga ttcctataga aagcattcct ggaattagag 2580
agacaggctg gaaaccgagt ggaaaagaga aaagtaaaga gccagagac cctgaccagc 2640
tttacagcac gctcaagagc atcctccagc aggtgaagag ccatcaaagc gcttggccct 2700
tcatggaacc tgtgaagaga acagaagtc caggatatta tgaagttata aggttcccca 2760
tggatctgaa aacctgagt gaacgcctca agaataggt ctacgtgtct aagaaattat 2820
tcatggcaga cttacagcga gtctttacca attgcaaaga gtacaacgcc gctgagagt 2880
aatactacaa atgtgccaat atcctggaga aattcttctt cagtaaaatt aaggaagctg 2940
gattaattga caagtgattt tttttcccc tctgcttctt agaaactcac caagcagtgt 3000
gcctaaagca aggt 3014

```

<210> 2

<211> 832

<212> PRT

<213> Homo sapiens

<400> 2

```

Met Ser Glu Ala Gly Gly Ala Gly Pro Gly Gly Cys Gly Ala Gly Ala
  1              5              10             15

```

```

Gly Ala Gly Ala Gly Pro Gly Ala Leu Pro Pro Gln Pro Ala Ala Leu
      20              25              30

```

```

Pro Pro Ala Pro Pro Gln Gly Ser Pro Cys Ala Ala Ala Ala Gly Gly
      35              40              45

```

```

Ser Gly Ala Cys Gly Pro Ala Thr Ala Val Ala Ala Ala Gly Thr Ala
      50              55              60

```

```

Glu Gly Pro Gly Gly Gly Gly Ser Ala Arg Ile Ala Val Lys Lys Ala
      65              70              75             80

```

```

Gln Leu Arg Ser Ala Pro Arg Ala Lys Lys Leu Glu Lys Leu Gly Val

```

85										90					95				
Tyr	Ser	Ala	Cys	Lys	Ala	Glu	Glu	Ser	Cys	Lys	Cys	Asn	Gly	Trp	Lys				
			100					105					110						
Asn	Pro	Asn	Pro	Ser	Pro	Thr	Pro	Pro	Arg	Ala	Asp	Leu	Gln	Gln	Ile				
		115					120					125							
Ile	Val	Ser	Leu	Thr	Glu	Ser	Cys	Arg	Ser	Cys	Ser	His	Ala	Leu	Ala				
	130					135					140								
Ala	His	Val	Ser	His	Leu	Glu	Asn	Val	Ser	Glu	Glu	Glu	Met	Asn	Arg				
145					150				155						160				
Leu	Leu	Gly	Ile	Val	Leu	Asp	Val	Glu	Tyr	Leu	Phe	Thr	Cys	Val	His				
			165					170						175					
Lys	Glu	Glu	Asp	Ala	Asp	Thr	Lys	Gln	Val	Tyr	Phe	Tyr	Leu	Phe	Lys				
			180					185					190						
Leu	Leu	Arg	Lys	Ser	Ile	Leu	Gln	Arg	Gly	Lys	Pro	Val	Val	Glu	Gly				
		195					200					205							
Ser	Leu	Glu	Lys	Lys	Pro	Pro	Phe	Glu	Lys	Pro	Ser	Ile	Glu	Gln	Gly				
	210					215					220								
Val	Asn	Asn	Phe	Val	Gln	Tyr	Lys	Phe	Ser	His	Leu	Pro	Ala	Lys	Glu				
225					230				235						240				
Arg	Gln	Thr	Ile	Val	Glu	Leu	Ala	Lys	Met	Phe	Leu	Asn	Arg	Ile	Asn				
			245					250						255					
Tyr	Trp	His	Leu	Glu	Ala	Pro	Ser	Gln	Arg	Arg	Leu	Arg	Ser	Pro	Asn				
		260						265					270						
Asp	Asp	Ile	Ser	Gly	Tyr	Lys	Glu	Asn	Tyr	Thr	Arg	Trp	Leu	Cys	Tyr				
	275						280					285							
Cys	Asn	Val	Pro	Gln	Phe	Cys	Asp	Ser	Leu	Pro	Arg	Tyr	Glu	Thr	Thr				
	290					295					300								
Gln	Val	Phe	Gly	Arg	Thr	Leu	Leu	Arg	Ser	Val	Phe	Thr	Val	Met	Arg				
305					310				315					320					
Arg	Gln	Leu	Leu	Glu	Gln	Ala	Arg	Gln	Glu	Lys	Asp	Lys	Leu	Pro	Leu				
			325					330					335						
Glu	Lys	Arg	Thr	Leu	Ile	Leu	Thr	His	Phe	Pro	Lys	Phe	Leu	Ser	Met				







<210> 3  
<211> 12  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: peptide

<220>  
<221> VARIANT  
<222> (2)  
<223> It represents 2 to 3 undesignated amino acids.  
They can be any amino acids.

<220>  
<221> VARIANT  
<222> (4)  
<223> It represents 5 to 8 undesignated amino acids.  
They can be any amino acids.

<220>  
<221> VARIANT  
<222> (6)  
<223> It represents one undesignated amino acid. It can  
be any amino acid.

<220>  
<221> VARIANT  
<222> (9)  
<223> It represents 5 undesignated amino acids. They can  
be any amino acids.

<220>  
<221> VARIANT  
<222> (5)  
<223> It can be any amino acid from the group of: P, K,  
or H.

<220>  
<221> VARIANT  
<222> (8)  
<223> It can be any amino acid from the group of: Y, F,  
or H.

<220>  
<221> VARIANT  
<222> (11)  
<223> It can be any amino acid from the group of: M, I,

or V.

<400> 3

Phe Xaa Pro Xaa Xaa Xaa Tyr Xaa Xaa Pro Xaa Asp  
1 5 10

<210> 4

<211> 12

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: peptide

<220>

<221> SITE

<222> (6)

<223> It is acetyl-lysine.

<400> 4

Ile Ser Tyr Gly Arg Xaa Lys Arg Arg Gln Arg Arg  
1 5 10

<210> 5

<211> 14

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: peptide

<220>

<221> SITE

<222> (8)

<223> It is acetyl-lysine.

<400> 5

Ala Arg Lys Ser Thr Gly Gly Xaa Ala Pro Arg Lys Gln Leu  
1 5 10

<210> 6

<211> 14

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: peptide

<220>

<221> SITE

<222> (8)

<223> It is acetyl-lysine.

<400> 6

Gln Ser Thr Ser Arg His Lys Xaa Leu Met Phe Lys Thr Glu  
1 5 10

<210> 7

<211> 110

<212> PRT

<213> Homo sapiens

<400> 7

Ser Lys Glu Pro Arg Asp Pro Asp Gln Leu Tyr Ser Thr Leu Lys Ser  
1 5 10 15

Ile Leu Gln Gln Val Lys Ser His Gln Ser Ala Trp Pro Phe Met Glu  
20 25 30

Pro Val Lys Arg Thr Glu Ala Pro Gly Tyr Tyr Glu Val Ile Arg Ser  
35 40 45

Pro Met Asp Leu Lys Thr Met Ser Glu Arg Leu Lys Asn Arg Tyr Tyr  
50 55 60

Val Ser Lys Lys Leu Phe Met Ala Asp Leu Gln Arg Val Phe Thr Asn  
65 70 75 80

Cys Lys Glu Tyr Asn Ala Pro Glu Ser Glu Tyr Tyr Lys Cys Ala Asn  
85 90 95

Ile Leu Glu Lys Phe Phe Phe Ser Lys Ile Lys Glu Ala Gly  
100 105 110

<210> 8

<211> 110

<212> PRT

<213> Homo sapiens

<400> 8

Gly Lys Glu Leu Lys Asp Pro Asp Gln Leu Tyr Thr Thr Leu Lys Asn



<210> 10  
 <211> 109  
 <212> PRT  
 <213> *Saccharomyces cerevisiae*

<400> 10  
 Ala Gln Arg Pro Lys Arg Gly Pro His Asp Ala Ala Ile Gln Asn Ile  
 1 5 10 15  
 Leu Thr Glu Leu Gln Asn His Ala Ala Ala Trp Pro Phe Leu Gln Pro  
 20 25 30  
 Val Asn Lys Glu Glu Val Pro Asp Tyr Tyr Asp Phe Ile Lys Glu Pro  
 35 40 45  
 Met Asp Leu Ser Thr Met Glu Ile Lys Leu Glu Ser Asn Lys Tyr Gln  
 50 55 60  
 Lys Met Glu Asp Phe Ile Tyr Asp Ala Arg Leu Val Phe Asn Asn Cys  
 65 70 75 80  
 Arg Met Tyr Asn Gly Glu Asn Thr Ser Tyr Tyr Lys Tyr Ala Asn Arg  
 85 90 95  
 Leu Glu Lys Phe Phe Asn Asn Lys Val Lys Glu Ile Pro  
 100 105

<210> 11  
 <211> 112  
 <212> PRT  
 <213> *Homo sapiens*

<400> 11  
 Lys Lys Ile Phe Lys Pro Glu Glu Leu Arg Gln Ala Leu Met Pro Thr  
 1 5 10 15  
 Leu Glu Ala Leu Tyr Arg Gln Asp Pro Glu Ser Leu Pro Phe Arg Gln  
 20 25 30  
 Pro Val Asp Pro Gln Leu Leu Gly Ile Pro Asp Tyr Phe Asp Ile Val  
 35 40 45  
 Lys Ser Pro Met Asp Leu Ser Thr Ile Lys Arg Lys Leu Asp Thr Gly  
 50 55 60  
 Gln Tyr Gln Glu Pro Trp Gln Tyr Val Asp Asp Ile Trp Leu Met Phe

0022220"4TET560

65		70		75		80
Asn Asn Ala Trp Leu Tyr Asn Arg Lys Thr Ser Arg Val Tyr Lys Tyr						
	85			90		95
Cys Ser Lys Leu Ser Glu Val Phe Glu Gln Glu Ile Asp Pro Val Met						
	100			105		110

<210> 12  
 <211> 112  
 <212> PRT  
 <213> Homo sapiens

<400> 12
Lys Lys Ile Phe Lys Pro Glu Glu Leu Arg Gln Ala Leu Met Pro Thr
1 5 10 15
Leu Glu Ala Leu Tyr Arg Gln Asp Pro Glu Ser Leu Pro Phe Arg Gln
20 25 30
Pro Val Asp Pro Gln Leu Leu Gly Ile Pro Asp Tyr Phe Asp Ile Val
35 40 45
Lys Asn Pro Met Asp Leu Ser Thr Ile Lys Arg Lys Leu Asp Thr Gly
50 55 60
Gln Tyr Gln Glu Pro Trp Gln Tyr Val Asp Asp Val Trp Leu Met Phe
65 70 75 80
Asn Asn Ala Trp Leu Tyr Asn Arg Lys Thr Ser Arg Val Tyr Lys Phe
85 90 95
Cys Ser Lys Leu Ala Glu Val Phe Glu Gln Glu Ile Asp Pro Val Met
100 105 110

<210> 13  
 <211> 112  
 <212> PRT  
 <213> Mus musculus



002220" TFE0T560

<400> 13

Lys	Lys	Ile	Phe	Lys	Pro	Glu	Glu	Leu	Arg	Gln	Ala	Leu	Met	Pro	Thr
1				5					10					15	
Leu	Glu	Ala	Leu	Tyr	Arg	Gln	Asp	Pro	Glu	Ser	Leu	Pro	Phe	Arg	Gln
			20						25				30		
Pro	Val	Asp	Pro	Gln	Leu	Leu	Gly	Ile	Pro	Asp	Tyr	Phe	Asp	Ile	Val
		35					40					45			
Lys	Asn	Pro	Met	Asp	Leu	Ser	Thr	Ile	Lys	Arg	Lys	Leu	Asp	Thr	Gly
	50					55					60				
Gln	Tyr	Gln	Glu	Pro	Trp	Gln	Tyr	Val	Asp	Asp	Val	Arg	Leu	Met	Phe
65					70					75					80
Asn	Asn	Ala	Trp	Leu	Tyr	Asn	Arg	Lys	Thr	Ser	Arg	Val	Tyr	Lys	Phe
				85					90					95	
Cys	Ser	Lys	Leu	Ala	Glu	Val	Phe	Glu	Gln	Glu	Ile	Asp	Pro	Val	Met
			100					105					110		

<210> 14

<211> 111

<212> PRT

<213> Caenorhabditis elegans

<400> 14

Asp	Thr	Val	Phe	Ser	Gln	Glu	Asp	Leu	Ile	Lys	Phe	Leu	Leu	Pro	Val
1				5					10					15	
Trp	Glu	Lys	Leu	Asp	Lys	Ser	Glu	Asp	Ala	Ala	Pro	Phe	Arg	Val	Pro
			20					25					30		
Val	Asp	Ala	Lys	Leu	Leu	Asn	Ile	Pro	Asp	Tyr	His	Glu	Ile	Ile	Lys
		35					40					45			
Arg	Pro	Met	Asp	Leu	Glu	Thr	Val	His	Lys	Lys	Leu	Tyr	Ala	Gly	Gln
	50					55					60				
Tyr	Gln	Asn	Ala	Gly	Gln	Phe	Cys	Asp	Asp	Ile	Trp	Leu	Met	Leu	Asp
65					70					75					80
Asn	Ala	Trp	Leu	Tyr	Asn	Arg	Lys	Asn	Ser	Lys	Val	Tyr	Lys	Tyr	Gly



Pro Met Asp Leu Gln Thr Leu Arg Glu Asn Val Arg Lys Arg Leu Tyr  
 50 55 60

Pro Ser Arg Glu Glu Phe Arg Glu His Leu Glu Leu Ile Val Lys Asn  
 65 70 75 80

Ser Ala Thr Tyr Asn Gly Pro Lys His Ser Leu Thr Gln Ile Ser Gln  
 85 90 95

Ser Met Leu Asp Leu Cys Asp Glu Lys Leu Lys Glu Lys Glu  
 100 105 110

<210> 17

<211> 111

<212> PRT

<213> Homo sapiens

<400> 17

Leu Leu Asp Asp Asp Gln Val Ala Phe Ser Phe Ile Leu Asp Asn  
 1 5 10 15

Ile Val Thr Gln Lys Met Met Ala Val Pro Asp Ser Trp Pro Phe His  
 20 25 30

His Pro Val Asn Lys Lys Phe Val Pro Asp Tyr Tyr Lys Val Ile Val  
 35 40 45

Asn Pro Met Asp Leu Glu Thr Ile Arg Lys Asn Ile Ser Lys His Lys  
 50 55 60

Tyr Gln Ser Arg Glu Ser Phe Leu Asp Asp Val Asn Leu Ile Leu Ala  
 65 70 75 80

Asn Ser Val Lys Tyr Asn Gly Pro Glu Ser Gln Tyr Thr Lys Thr Ala  
 85 90 95

Gln Glu Ile Val Asn Val Cys Tyr Gln Thr Leu Thr Glu Tyr Asp  
 100 105 110

<210> 18

<211> 111

<212> PRT

<213> Mesocricetus auratus

<400> 18





Tyr Trp Ser Ala Lys Glu Thr Ile Gln Asp Phe Asn Thr Met Phe Asn  
65 70 75 80

Asn Cys Tyr Val Tyr Asn Lys Pro Gly Glu Asp Val Val Val Met Ala  
85 90 95

Gln Thr Leu Glu Lys Val Phe Leu Gln Lys Ile Glu Ser Met Pro  
100 105 110

<210> 22

<211> 109

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 22

Asn Pro Ile Pro Lys His Gln Gln Lys His Ala Leu Leu Ala Ile Lys  
1 5 10 15

Ala Val Lys Arg Leu Lys Asp Ala Arg Pro Phe Leu Gln Pro Val Asp  
20 25 30

Pro Val Lys Leu Asp Ile Pro Phe Tyr Phe Asn Tyr Ile Lys Arg Pro  
35 40 45

Met Asp Leu Ser Thr Ile Glu Arg Lys Leu Asn Val Gly Ala Tyr Glu  
50 55 60

Val Pro Glu Gln Ile Thr Glu Asp Phe Asn Leu Met Val Asn Asn Ser  
65 70 75 80

Ile Lys Phe Asn Gly Pro Asn Ala Gly Ile Ser Gln Met Ala Arg Asn  
85 90 95

Ile Gln Ala Ser Phe Glu Lys His Met Leu Asn Met Pro  
100 105

<210> 23

<211> 113

<212> PRT

<213> *Homo sapiens*

<400> 23

Lys Lys Gly Lys Leu Ser Glu Gln Leu Lys His Cys Asn Gly Ile Leu  
1 5 10 15

Lys Glu Leu Leu Ser Lys Lys His Ala Ala Tyr Ala Trp Pro Phe Tyr

002220" 4 FEB 1959

	20							25								30
Lys	Pro	Val	Asp	Ala	Ser	Ala	Leu	Gly	Leu	His	Asp	Tyr	His	Asp	Ile	
	35						40					45				
Ile	Lys	His	Pro	Met	Asp	Leu	Ser	Thr	Val	Lys	Arg	Lys	Met	Glu	Asn	
	50					55					60					
Arg	Asp	Tyr	Arg	Asp	Ala	Gln	Glu	Phe	Ala	Ala	Asp	Val	Arg	Leu	Met	
65					70					75					80	
Phe	Ser	Asn	Cys	Tyr	Lys	Tyr	Asn	Pro	Pro	Asp	His	Asp	Val	Val	Ala	
				85					90					95		
Met	Ala	Arg	Lys	Leu	Gln	Asp	Val	Phe	Glu	Phe	Arg	Tyr	Ala	Lys	Met	
		100						105					110			

Pro

<210> 24  
<211> 113  
<212> PRT  
<213> Homo sapiens

<400>	24															
Lys	Lys	Gly	Lys	Leu	Ser	Glu	His	Leu	Arg	Tyr	Cys	Asp	Ser	Ile	Leu	
1				5					10					15		
Arg	Glu	Met	Leu	Ser	Lys	Lys	His	Ala	Ala	Tyr	Ala	Trp	Pro	Phe	Tyr	
			20					25					30			
Lys	Pro	Val	Asp	Ala	Glu	Ala	Leu	Glu	Leu	His	Asp	Tyr	His	Asp	Ile	
	35						40					45				
Ile	Lys	His	Pro	Met	Asp	Leu	Ser	Thr	Val	Lys	Arg	Lys	Met	Asp	Gly	
	50					55					60					
Arg	Glu	Tyr	Pro	Asp	Ala	Gln	Gly	Phe	Ala	Ala	Asp	Val	Arg	Leu	Met	
65					70					75					80	
Phe	Ser	Asn	Cys	Tyr	Lys	Tyr	Asn	Pro	Pro	Asp	His	Glu	Val	Val	Ala	
				85					90					95		
Met	Ala	Arg	Lys	Leu	Gln	Asp	Val	Phe	Glu	Met	Arg	Phe	Ala	Lys	Met	
		100						105					110			

Pro

<210> 25

<211> 113

<212> PRT

<213> *Drosophila melanogaster*

<400> 25

Asn Lys Glu Lys Leu Ser Asp Ala Leu Lys Ser Cys Asn Glu Ile Leu  
1 5 10 15

Lys Glu Leu Phe Ser Lys Lys His Ser Gly Tyr Ala Trp Pro Phe Tyr  
20 25 30

Lys Pro Val Asp Ala Glu Met Leu Gly Leu His Asp Tyr His Asp Ile  
35 40 45

Ile Lys Lys Pro Met Asp Leu Gly Thr Val Lys Arg Lys Met Asp Asn  
50 55 60

Arg Glu Tyr Lys Ser Ala Pro Glu Phe Ala Ala Asp Val Arg Leu Ile  
65 70 75 80

Phe Thr Asn Cys Tyr Lys Tyr Asn Pro Pro Asp His Asp Val Val Ala  
85 90 95

Met Gly Arg Lys Leu Gln Asp Val Phe Glu Met Arg Tyr Ala Asn Ile  
100 105 110

Pro

<210> 26

<211> 113

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 26

Lys Ser Lys Arg Leu Gln Gln Ala Met Lys Phe Cys Gln Ser Val Leu  
1 5 10 15

Lys Glu Leu Met Ala Lys Lys His Ala Ser Tyr Asn Tyr Pro Phe Leu  
20 25 30

Glu Pro Val Asp Pro Val Ser Met Asn Leu Pro Thr Tyr Phe Asp Tyr



35                      40                      45  
 Val Lys Glu Pro Met Asp Leu Gly Thr Ile Ala Lys Lys Leu Asn Asp  
     50                      55                      60  
 Trp Gln Tyr Gln Thr Met Glu Asp Phe Glu Arg Glu Val Arg Leu Val  
     65                      70                      75                      80  
 Phe Lys Asn Cys Tyr Thr Phe Asn Pro Asp Gly Thr Ile Val Asn Met  
                     85                      90                      95  
 Met Gly His Arg Leu Glu Glu Val Phe Asn Ser Lys Trp Ala Asp Arg  
                     100                      105                      110

Pro

<210> 27  
 <211> 108  
 <212> PRT  
 <213> Homo sapiens

<400> 27  
 Met Glu Met Gln Leu Thr Pro Phe Leu Ile Leu Leu Arg Lys Thr Leu  
     1                      5                      10                      15  
 Glu Gln Leu Gln Glu Lys Asp Thr Gly Asn Ile Phe Ser Glu Pro Val  
                     20                      25                      30  
 Pro Leu Ser Glu Val Pro Asp Tyr Leu Asp His Ile Lys Lys Pro Met  
                     35                      40                      45  
 Asp Phe Phe Thr Met Lys Gln Asn Leu Glu Ala Tyr Arg Tyr Leu Asn  
                     50                      55                      60  
 Phe Asp Asp Phe Glu Glu Asp Phe Asn Leu Ile Val Ser Asn Cys Leu  
     65                      70                      75                      80  
 Lys Tyr Asn Ala Lys Asp Thr Ile Phe Tyr Arg Ala Ala Val Arg Leu  
                     85                      90                      95  
 Arg Glu Gln Gly Gly Ala Val Val Arg Gln Ala Arg  
                     100                      105

<210> 28  
 <211> 113

<212> PRT  
<213> Homo sapiens

<400> 28

Ser Glu Asp Gln Glu Ala Ile Gln Ala Gln Lys Ile Trp Lys Lys Ala  
1 5 10 15

Ile Met Leu Val Trp Arg Ala Ala Ala Asn His Arg Tyr Ala Asn Val  
20 25 30

Phe Leu Gln Pro Val Thr Asp Asp Ile Ala Pro Gly Tyr His Ser Ile  
35 40 45

Val Gln Arg Pro Met Asp Leu Ser Thr Ile Lys Lys Asn Ile Glu Asn  
50 55 60

Gly Leu Ile Arg Ser Thr Ala Glu Phe Gln Arg Asp Ile Met Leu Met  
65 70 75 80

Phe Gln Asn Ala Val Met Tyr Asn Ser Ser Asp His Asp Val Tyr His  
85 90 95

Met Ala Val Glu Met Gln Arg Asp Val Leu Glu Gln Ile Gln Gln Phe  
100 105 110

Leu

<210> 29  
<211> 106  
<212> PRT  
<213> Gallus gallus

<400> 29

Asn Leu Pro Thr Val Asp Pro Ile Ala Val Cys His Glu Leu Tyr Asn  
1 5 10 15

Thr Ile Arg Asp Tyr Lys Asp Glu Gln Gly Arg Leu Leu Cys Glu Leu  
20 25 30

Phe Ile Arg Ala Pro Lys Arg Arg Asn Gln Pro Asp Tyr Tyr Glu Val  
35 40 45

Val Ser Gln Pro Ile Asp Leu Met Lys Ile Gln Gln Lys Leu Lys Met  
50 55 60

Glu Glu Tyr Asp Asp Val Asn Val Leu Thr Ala Asp Phe Gln Leu Leu

65 70 75 80

Phe Asn Asn Ala Lys Ala Tyr Tyr Lys Pro Asp Ser Pro Glu Tyr Lys  
85 90 95

Ala Ala Cys Lys Leu Trp Glu Leu Tyr Leu  
100 105

<210> 30

<211> 112

<212> PRT

<213> Gallus gallus

<400> 30

Ser Ser Pro Gly Tyr Leu Lys Glu Ile Leu Glu Gln Leu Leu Glu Ala  
1 5 10 15

Val Ala Val Ala Thr Asn Pro Ser Gly Arg Leu Ile Ser Glu Leu Phe  
20 25 30

Gln Lys Leu Pro Ser Lys Val Gln Tyr Pro Asp Tyr Tyr Ala Ile Ile  
35 40 45

Lys Glu Pro Ile Asp Leu Lys Thr Ile Ala Gln Arg Ile Gln Asn Gly  
50 55 60

Thr Tyr Lys Ser Ile His Ala Met Ala Lys Asp Ile Asp Leu Leu Ala  
65 70 75 80

Lys Asn Ala Lys Thr Tyr Asn Glu Pro Gly Ser Gln Val Phe Lys Asp  
85 90 95

Ala Asn Ala Ile Lys Lys Ile Phe Asn Met Lys Lys Ala Glu Ile Glu  
100 105 110

<210> 31

<211> 112

<212> PRT

<213> Gallus gallus

<400> 31

Thr Ser Phe Met Asp Thr Ser Asn Pro Leu Tyr Gln Leu Tyr Asp Thr  
1 5 10 15

Val	Arg	Ser	Cys	Arg	Asn	Asn	Gln	Gly	Gln	Leu	Ile	Ser	Glu	Pro	Phe
			20					25					30		
Phe	Gln	Leu	Pro	Ser	Lys	Lys	Lys	Tyr	Pro	Asp	Tyr	Tyr	Gln	Gln	Ile
		35					40					45			
Lys	Thr	Pro	Ile	Ser	Leu	Gln	Gln	Ile	Arg	Ala	Lys	Leu	Lys	Asn	His
	50					55					60				
Glu	Tyr	Glu	Thr	Leu	Asp	Gln	Leu	Glu	Ala	Asp	Leu	Asn	Leu	Met	Phe
	65				70					75					80
Glu	Asn	Ala	Lys	Arg	Tyr	Asn	Val	Pro	Asn	Ser	Ala	Ile	Tyr	Lys	Arg
				85					90					95	
Val	Leu	Lys	Met	Gln	Gln	Val	Met	Gln	Ala	Lys	Lys	Lys	Glu	Leu	Ala
			100					105					110		

<210> 32  
 <211> 113  
 <212> PRT  
 <213> Gallus gallus

<400> 32															
Ser	Lys	Lys	Asn	Met	Arg	Lys	Gln	Arg	Met	Lys	Ile	Leu	Tyr	Asn	Ala
1				5					10					15	
Val	Leu	Glu	Ala	Arg	Glu	Ser	Gly	Thr	Gln	Arg	Arg	Leu	Cys	Asp	Leu
			20					25					30		
Phe	Met	Val	Lys	Pro	Ser	Lys	Lys	Asp	Tyr	Pro	Asp	Tyr	Tyr	Lys	Ile
		35					40					45			
Ile	Leu	Glu	Pro	Met	Asp	Leu	Lys	Met	Ile	Glu	His	Asn	Ile	Arg	Asn
	50						55				60				
Asp	Lys	Tyr	Val	Gly	Glu	Glu	Ala	Met	Ile	Asp	Asp	Met	Lys	Leu	Met
	65				70					75					80
Phe	Arg	Asn	Ala	Arg	His	Tyr	Asn	Glu	Glu	Gly	Ser	Gln	Val	Tyr	Asn
				85				90					95		
Asp	Ala	His	Met	Leu	Glu	Lys	Ile	Leu	Lys	Glu	Lys	Arg	Lys	Glu	Leu

100 105 110

Gly

<210> 33  
 <211> 115  
 <212> PRT  
 <213> Gallus gallus

<400> 33  
 Lys Lys Ser Lys Tyr Met Thr Pro Met Gln Gln Lys Leu Asn Glu Val  
 1 5 10 15  
 Tyr Glu Ala Val Lys Asn Tyr Thr Asp Lys Arg Gly Arg Arg Leu Ser  
 20 25 30  
 Ala Ile Phe Leu Arg Leu Pro Ser Arg Ser Glu Leu Pro Asp Tyr Tyr  
 35 40 45  
 Ile Thr Ile Lys Lys Pro Val Asp Met Glu Lys Ile Arg Ser His Met  
 50 55 60  
 Met Ala Asn Lys Tyr Gln Asp Ile Asp Ser Met Val Glu Asp Phe Val  
 65 70 75 80  
 Met Met Phe Asn Asn Ala Cys Thr Tyr Asn Glu Pro Glu Ser Leu Ile  
 85 90 95  
 Tyr Lys Asp Ala Leu Val Leu His Lys Val Leu Leu Glu Thr Arg Arg  
 100 105 110  
 Glu Ile Glu  
 115

<210> 34  
 <211> 112  
 <212> PRT  
 <213> Unknown

<220>  
 <223> Description of Unknown Organism: Cited from  
 Jeanmougin et al., Trends in Biochemical Sciences,  
 22:151-153 (1997)

<400> 34

His Asn Ala Pro Phe Asp Lys Thr Lys Phe Asp Glu Val Leu Glu Ala  
1 5 10 15

Leu Val Gly Leu Lys Asp Asn Glu Gly Asn Pro Phe Asp Asp Ile Phe  
20 25 30

Glu Glu Leu Pro Ser Lys Arg Tyr Phe Pro Asp Tyr Tyr Gln Ile Ile  
35 40 45

Gln Lys Pro Ile Cys Tyr Lys Met Met Arg Asn Lys Ala Lys Thr Gly  
50 55 60

Lys Tyr Leu Ser Met Gly Asp Phe Tyr Asp Asp Ile Arg Leu Met Val  
65 70 75 80

Ser Asn Ala Gln Thr Tyr Asn Met Pro Gly Ser Leu Val Tyr Glu Cys  
85 90 95

Ser Val Leu Ile Ala Asn Thr Ala Asn Ser Leu Glu Ser Lys Asp Gly  
100 105 110

<210> 35

<211> 113

<212> PRT

<213> Unknown

<220>

<223> Description of Unknown Organism: Cited from  
Jeanmougin et al., Trends in Biochemical Sciences,  
22:151-153 (1997)

<400> 35

Gly Thr Asn Glu Ile Asp Val Pro Lys Val Ile Gln Asn Ile Leu Asp  
1 5 10 15

Ala Leu His Glu Glu Lys Asp Glu Gln Gly Arg Phe Leu Ile Asp Ile  
20 25 30

Phe Ile Asp Leu Pro Ser Lys Arg Leu Tyr Pro Asp Tyr Tyr Glu Ile  
35 40 45

Ile Lys Ser Pro Met Thr Ile Lys Met Leu Glu Lys Arg Phe Lys Lys  
50 55 60



<400> 37

Ser Pro Asn Pro Pro Asn Leu Thr Lys Lys Met Lys Lys Ile Val Asp  
1 5 10 15

Ala Val Ile Lys Tyr Lys Asp Ser Ser Ser Gly Arg Gln Leu Ser Glu  
20 25 30

Val Phe Ile Gln Leu Pro Ser Arg Lys Glu Leu Pro Glu Tyr Tyr Glu  
35 40 45

Leu Ile Arg Lys Pro Val Asp Phe Lys Lys Ile Lys Glu Arg Ile Arg  
50 55 60

Asn His Lys Tyr Arg Ser Leu Asn Asp Leu Glu Lys Asp Val Met Leu  
65 70 75 80

Leu Cys Gln Asn Ala Gln Thr Phe Asn Leu Glu Gly Ser Leu Ile Tyr  
85 90 95

Glu Asp Ser Ile Val Leu Gln Ser Val Phe Thr Ser Val Arg Gln Lys  
100 105 110

Ile Glu

<210> 38

<211> 113

<212> PRT

<213> Gallus gallus

<400> 38

Ser Pro Asn Pro Pro Lys Leu Thr Lys Gln Met Asn Ala Ile Ile Asp  
1 5 10 15

Thr Val Ile Asn Tyr Lys Asp Ser Ser Gly Arg Gln Leu Ser Glu Val  
20 25 30

Phe Ile Gln Leu Pro Ser Arg Lys Glu Leu Pro Glu Tyr Tyr Glu Leu  
35 40 45

Ile Arg Lys Pro Val Asp Phe Lys Lys Ile Lys Glu Arg Ile Arg Asn  
50 55 60

His Lys Tyr Arg Ser Leu Gly Asp Leu Glu Lys Asp Val Met Leu Leu  
65 70 75 80



Cys His Asn Ala Gln Thr Phe Asn Leu Glu Gly Ser Gln Ile Tyr Glu  
85 90 95

Asp Ser Ile Val Leu Gln Ser Val Phe Lys Ser Ala Arg Gln Lys Ile  
100 105 110

Ala

<210> 39  
<211> 114  
<212> PRT  
<213> Gallus gallus

<400> 39  
Ser Pro Asn Pro Pro Asn Leu Thr Lys Lys Met Lys Lys Ile Val Asp  
1 5 10 15

Ala Val Ile Lys Tyr Lys Asp Ser Ser Ser Gly Arg Gln Leu Ser Glu  
20 25 30

Val Phe Ile Gln Leu Pro Ser Arg Lys Glu Leu Pro Glu Tyr Tyr Glu  
35 40 45

Leu Ile Arg Lys Pro Val Asp Phe Lys Lys Ile Lys Glu Arg Ile Arg  
50 55 60

Asn His Lys Tyr Arg Ser Leu Asn Asp Leu Glu Lys Asp Val Met Leu  
65 70 75 80

Leu Cys Gln Asn Ala Gln Thr Phe Asn Leu Glu Val Ser Leu Ile Tyr  
85 90 95

Glu Asp Ser Ile Val Leu Gln Ser Val Phe Thr Ser Val Arg Gln Lys  
100 105 110

Ile Glu

<210> 40  
<211> 105  
<212> PRT  
<213> Homo sapiens

<400> 40  
Ala Lys Leu Ser Pro Ala Asn Gln Arg Lys Cys Glu Arg Val Leu Leu

1	5	10	15
Ala Leu Phe Cys His Glu Pro Cys Arg Pro Leu His Gln Leu Ala Thr	20	25	30
Asp Ser Thr Phe Ser Leu Asp Gln Pro Gly Gly Thr Leu Asp Leu Thr	35	40	45
Leu Ile Arg Ala Arg Leu Gln Glu Lys Leu Ser Pro Pro Tyr Ser Ser	50	55	60
Pro Gln Glu Phe Ala Gln Asp Val Gly Arg Met Phe Lys Gln Phe Asn	65	70	75
Lys Leu Thr Glu Asp Lys Ala Asp Val Gln Ser Ile Ile Gly Leu Gln	85	90	95
Arg Phe Phe Glu Thr Arg Met Asn Glu	100	105	

<210> 41  
 <211> 105  
 <212> PRT  
 <213> Mus musculus

<400> 41

Ala Lys Leu Ser Pro Ala Asn Gln Arg Lys Cys Glu Arg Val Leu Leu	1	5	10	15
Ala Leu Phe Cys His Glu Pro Cys Arg Pro Leu His Gln Leu Ala Thr	20	25	30	
Asp Ser Thr Phe Ser Met Glu Gln Pro Gly Gly Thr Leu Asp Leu Thr	35	40	45	
Leu Ile Arg Ala Arg Leu Gln Glu Lys Leu Ser Pro Pro Tyr Ser Ser	50	55	60	
Pro Gln Glu Phe Ala Gln Asp Val Gly Arg Met Phe Lys Gln Phe Asn	65	70	75	80
Lys Leu Thr Glu Asp Lys Ala Asp Val Gln Ser Ile Ile Gly Leu Gln	85	90	95	
Arg Phe Phe Glu Thr Arg Met Asn Asp	100	105		

<210> 42  
 <211> 108  
 <212> PRT  
 <213> Mus sp.

<400> 42  
 Thr Lys Leu Thr Pro Ile Asp Lys Arg Lys Cys Glu Arg Leu Leu Leu  
 1 5 10 15  
 Phe Leu Tyr Cys His Glu Met Ser Leu Ala Phe Gln Asp Pro Val Pro  
 20 25 30  
 Leu Thr Val Pro Asp Tyr Tyr Lys Ile Ile Lys Asn Pro Met Asp Leu  
 35 40 45  
 Ser Thr Ile Lys Lys Arg Leu Gln Glu Asp Tyr Cys Met Tyr Thr Lys  
 50 55 60  
 Pro Glu Asp Phe Val Ala Asp Phe Arg Leu Ile Phe Gln Asn Cys Ala  
 65 70 75 80  
 Glu Phe Asn Glu Pro Asp Ser Glu Val Ala Asn Ala Gly Ile Lys Leu  
 85 90 95  
 Glu Ser Tyr Phe Glu Glu Leu Leu Lys Asn Leu Tyr  
 100 105

<210> 43  
 <211> 13  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: consencus

<220>  
 <221> VARIANT  
 <222> (1)  
 <223> It represents 2 amino acids. They can be any amino acids.

<220>  
 <221> VARIANT  
 <222> (3)  
 <223> It represents 2 to 3 amino acids. They can be any amino acids.

<220>  
 <221> VARIANT  
 <222> (5)  
 <223> It represents 5 to 8 amino acids. They can be any amino acids.

<220>  
 <221> VARIANT  
 <222> (7)  
 <223> It represents one amino acids. It can be any amino acid.

<220>  
 <221> VARIANT  
 <222> (10)  
 <223> It represents 5 amino acids. They can be any amino acids.

<220>  
 <221> VARIANT  
 <222> (6)  
 <223> It represents any amino acid from the group of: P, K, or H.

<220>  
 <221> VARIANT  
 <222> (9)  
 <223> It represents any amino acid from the group of: Y, F, or H.

<220>  
 <221> VARIANT  
 <222> (12)  
 <223> It represents any amino acid from the group of: M, I, or V.

<400> 43  
 Xaa Phe Xaa Pro Xaa Xaa Xaa Tyr Xaa Xaa Pro Xaa Asp  
   1                  5                  10

<210> 44  
 <211> 20  
 <212> PRT  
 <213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: consencus

<400> 44

Trp	Pro	Phe	Met	Glu	Pro	Val	Lys	Arg	Thr	Glu	Ala	Pro	Gly	Tyr	Tyr
1				5					10					15	

Glu	Val	Ile	Arg
			20

002220"4TE07560